GRAMMATICAL NUMBER AND THE SCALE OF INDIVIDUATION

Scott Grimm

University of Rochester

A central debate in the literature on grammatical number systems and nominal semantics is whether the countable/noncountable contrast is ontologically based or ultimately arbitrary. This article examines this question in light of several languages that express three or more categories of grammatical number, in particular including a collective category containing nouns of an intermediate status between prototypical countable nouns and prototypical noncountable nouns. I connect this crosslinguistic data to psycholinguistic research on INDIVIDUATION, identifying several INDIVIDUATION TYPES, that is, noun meanings organized into equivalence classes based on shared individuation properties. The individuation types themselves can be ordered, giving rise to a SCALE OF INDIVIDUATION. I propose that the organization of grammatical number systems reflects the scale of individuation, effectively steering a middle course between ontological and grammatical accounts. This approach accounts for a range of grammatical number systems and makes broad predictions bearing on what possible grammatical number systems are.*

Keywords: grammatical number, mass/count, collectives, individuation, animacy

1. INTRODUCTION. The noun *dog* uses morphological and syntactic means to be counted, such as pluralization, but the noun *water* does not. Average speakers of English are hardly mystified by this contrast—for *dog* names a type of object in the world that is easily countable, namely a discrete entity, while *water* refers to a (nondiscrete) substance. Despite the strong intuition that grammatical form and ontological type are related, this putative relation is the central point of controversy in the count/mass literature.

Many researchers have proposed that semantic categorization does not underlie the count/mass distinction. One argument hinges on crosslinguistic incommensurability: if the count/mass distinction reflects the ontological nature of entities in the world, namely discrete entities versus substances, it is unexpected that a particular entity would be named by a countable noun in one language but a noncountable noun in another. Yet such doublets are easy to find (see, inter alia, Palmer 1971, Ware 1975, Chierchia 1998, Pelletier & Schubert 2004, Rothstein 2010). For instance, *hair* in English permits noncountable syntax, while its French equivalent, *cheveux*, is obligatorily plural (Palmer 1971). A second argument concerns nouns, like *leaves* and *foliage*, that on certain occasions may describe the same stuff, yet nonetheless fall into different countability classes, countable and noncountable, respectively (Palmer 1971). Again, so the

* I would like to thank Greg Carlson, Beth Levin, Louise McNally, three anonymous referees, and associate editor Chris Kennedy for helpful comments on drafts of this article. This project grew out of trying to understand the inverse number-marking system of Dagaare, and I would like to thank Adams Bodomo for his early and continued encouragement. I would also like to express my profound thanks to Mark Ali at the University of Education in Winneba, Ghana, who devoted an enormous amount of time to working with me on Dagaare. I would also like to thank the many others who provided helpful input on this project, particularly Arto Anttila, Jade Comfort, Henriëtte de Swart, Alex Djalali, Matthew Dryer, Donka Farkas, Hana Filip, Jeff Good, Paul Kiparksy, Sven Lauer, Beth Levin, John Lucy, Chris Potts, Tania Rojas-Esponda, Frank Seifart, Tom Wasow, and Henk Zeevat; as well as the audiences at the Department of African Studies at Humboldt University, ILLC Amsterdam, MPI Leipzig, MPI Nijmegen, UC Berkeley, University of Buffalo, University of Rochester, University of Utrecht, ZAS, and at the Workshop on Countability at Heinrich-Heine-University. This material is in part based upon work supported by the National Science Foundation under Grant No. 1023289 'Number and individuation: Nominal semantics in Dagaare'. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Printed with the permission of Scott Grimm. © 2018.

argument goes, if the grammatical count/mass distinction is a reflection of an ontological distinction, such interchangeability is unexpected.

Based on such evidence, many have come to the same conclusion as found in Chierchia 1998:57: 'while the mass/count distinction is not altogether indifferent to how things are inherently structured, it appears to be independent of it, which is what makes such a distinction a strictly grammatical one'. This critique targets a semantic theory of countability where there is a direct, one-to-one correspondence between, for instance, discrete entities in the world and countable nouns, and substances and noncountable nouns. These authors are correct that a strong view upon which the count/mass distinction directly reflects a binary ontological distinction between objects and substances faces fatal challenges from crosslinguistic and intralinguistic data. Yet two important assumptions underlie this argument: (i) countability is a binary phenomenon (count versus mass) and (ii) the ONLY relevant extralinguistic distinction is the object/substance distinction. I argue that neither assumption holds.

First, many languages possess three or more categories of grammatical number, indicating that countability is richer than a binary count/mass contrast. Second, increasing evidence from the psycholinguistic literature has shown that, while the object/substance distinction is important, many other factors also play a role, such as ease of distinguishing individual units. Taken together, this indicates that grammatical countability is a much more heterogeneous domain than typically assumed, with multiple grammatical categories of number and multiple extralinguistic factors that influence into which of those categories a noun will fall.

This article further rejects the supposition that countability must be either a grammatical phenomenon or a semantic one; it instead steers a middle course, recognizing that both semantic and grammatical categorization are active in the organization of grammatical number systems. I argue that while there is certainly some variation as to how grammatical number systems are organized, they also obey broad semantic constraints.

The article examines the interplay between COUNTABILITY and INDIVIDUATION. I use the term *countability* to refer to morphosyntactic distinctions made in languages related to grammatical number, such as singular, plural, collective, and so forth, and the grammatical categories formed upon these distinctions, for instance, the class of nouns with a singular/plural contrast. Section 2 contributes data from several languages that make more than a binary countability contrast. I investigate in particular these languages' collective classes, which have an intermediate status between prototypical countable nouns and prototypical noncountable nouns.

Individuation refers to the propensity for an entity to be construed as an independent individual.[1] After reviewing relevant results from the psycholinguistic literature in §3.1, I argue that individuation is not a binary distinction (individual versus nonindividual), but a scalar distinction, which tracks the degree to which the referents of nouns are construed as independent individuals. I then argue that several individuation types can be identified, wherein noun meanings are organized into equivalence classes based on shared individuation properties, and the individuation types can be ordered based on the degree of individuation, giving rise to a scale of individuation.

The core thesis of the article is that the organization of grammatical number systems reflects the organization present in the scale of individuation. This framework yields a

[1] Another use of the term *individuation* in the philosophical literature relates to referentially individuating objects, and in particular to the sortal/nonsortal distinction and problems of identity (Wiggins 1980, Brennan 1988). Although countability is ultimately also bound up with referential aspects of language use, I do not pursue those connections here.

range of predictions, such as what possible grammatical number systems are and what form they may take, that is, patterns of markedness in grammatical number systems. In §4, I explore the typological predictions of this scalar view on individuation and consider how this view connects with other factors known to be relevant, such as animacy.

Section 5 discusses how this view speaks to both the regularity and variation found in grammatical number systems. Grounding countability distinctions in entities in the world and our perception of and interaction with them provides an explanation for the regularity found across grammatical countability systems, since the physical properties as well as how speech communities perceive or interact with, say, water are likely to be quite commensurable. Yet the model does not overdetermine how reality is mapped to vocabulary: there is often a choice to be made in how entities are to be described and which perceptual or interactional properties of an entity are highlighted. There is a further decision point concerning how many and what type of grammatical categories a language manifests, but this is constrained by the scale of individuation. This view then allows for principled variation, and further provides predictions as to where variability is more likely to (not) be found and what forms of variability one might expect. I show how this view extends to treat cases of lexical variation (*leaves* versus *foliage*) as well as contextual variation due to 'grinding' and 'packaging' uses. I then conclude in §6.

Before proceeding, it is useful to fix some terminology that I use throughout the rest of the article. While the subject of this article is normally described as a distinction between 'count' and 'mass', the various ways these terms have been used have rendered them unfortunately confusing. For some, *mass* is taken to designate a grammatical category (Bloomfield 1933 and many following him), but for others, it has a narrower range and aligns only with certain types of entities in the world, namely substances (see discussion in Joosten 2003). I instead describe nouns as COUNTABLE or NONCOUNTABLE, where these two terms are restricted to designating nominal behavior in terms of morphosyntactic characteristics and make no reference to semantic characteristics. I also draw a terminological distinction between ENTITY, referring to some portion or element of the external world, and INDIVIDUAL, referring to a construal of an entity as an individual object.

2. BEYOND BINARY SYSTEMS. A set of distributional characteristics have been taken as reflecting a distinction between countable and noncountable nouns in English. As summarized in Table 1, countable nouns, like *dog*, pluralize (*dogs*) and combine with numerals (*two dogs*), determiners implicating cardinality (*several dogs*), and the indefinite article (*a dog*). Noncountable nouns permit neither pluralization nor these other combinations. Instead, noncountable nouns may appear bare or combine with quantifiers such as *much* and *a little* or with measure terms such as *three kilos of*.

MORPHOSYNTACTIC TRAITS	SINGULARS	PLURALS	NONCOUNTABLE
pluralization	✓	N/A	*
quantifiers implicating plurality (*many*, *several*)	*	✓	*
cardinal modification (*two*)	*	✓	*
much, *little*	*	*	✓
combine with measure terms (*two kilos*/*meters of*)	*	✓	✓

TABLE 1. Core distributional properties of nominal countability.

The picture of countability presented in Table 1 leads to the false hope that one could, thanks to these distributional properties, sort the lexicon into nouns that are countable and nouns that are noncountable. The situation is, however, much more complicated. First, many nouns may have several uses, some countable and some not, so these diagnostics

must be used with great care (see discussion in Pelletier & Schubert 2004). Second, not all nouns that could be classified as countable or noncountable by SOME distributional properties behave identically with respect to ALL of the distributional properties. As discussed in Allan 1980, there are many more nuanced distinctions than a strictly binary division can represent; for example, the noun *admiration* combines with the indefinite article (*a great admiration*), but does not combine with quantifiers implicating plurality (**several admirations*). Unsurprisingly, different grammatical number systems also make distinctions that do not completely align with a simple division into countable and noncountable nouns.

This section examines grammatical number systems of Welsh, Turkana, Maltese, and Dagaare, which express a greater number of countability distinctions beyond the simple binary countable/noncountable contrast expressed in English. For each language, I discuss how the countable/noncountable distinction is manifested and then other number distinctions that the language makes.

I have restricted this investigation in several ways. First, I focus on languages that code number distinctions through affixes that are unambiguously devoted to coding those distinctions. I do not discuss in detail languages with systems of classifiers or elaborate gender systems, since a possible confound lurks in those systems' tendency to code number along with other categories of nominal meaning, such as shape or animacy (see Denny 1976, Maho 1999, Katamba 2003 for Niger-Congo systems).

Second, I focus on what I term NATURAL CONCRETE ENTITIES (*dog*, *water*) to the exclusion of ARTIFACTUAL ENTITIES (*hammer*, *furniture*) or ABSTRACT ENTITIES (*arrival*, *happiness*). Natural concrete entities provide the firmest foundation for comparative study since the relation between nouns and their referents is reasonably clear. If two languages have nouns for describing the entity 'apple', it is reasonable to assume that these nouns refer to the same thing. Less clear is whether words glossed as 'sadness' across different languages refer to the same (abstract) entity. Cultures differ widely as to how they describe and lexicalize abstract notions, such as emotions, and this wide variation obstructs comparison across languages. Further, artifactual and abstract nouns, as the examples given indicate, tend to be derived nouns, and this adds another layer of complexity that would hinder the investigation. I return to the issue of artifactual and abstract nouns in §4.3.

Finally, while it is well known that a particular noun's countability status may be able to shift due to contextual influence, in this section and in §3 I focus on comparing canonical uses of nouns across languages. Thus, while *dog* may have a use in English that designates dog-stuff in GRINDING contexts such as *dog all over the street*, I focus on its canonical use as designating a living being, and similarly for PACKAGING uses of nouns like *water*, such as *the waiter brought three waters*. The primary reason to set them aside is that there is a high degree of variation, both as to which nouns within a language permit multiple uses, and as to which languages allow different types of nominal flexibility. To illustrate the last point, languages considered here include those that are highly permissive of packaging (Yudja) and also those for which packaging is virtually impossible (Dagaare). I return to contextual variation in §5.1, but for now, I focus on understanding the core organization of grammatical number systems and countability.

2.1. COUNTABILITY IN WELSH. Welsh resembles English in possessing certain canonical morphosyntactic diagnostics of countability, which distinguish, for example, noncountable nouns. Yet Welsh differs from English in possessing—in addition to nouns where the singular value is zero-coded and the plural is overtly coded—a class of nouns

for which the plural value is zero-coded (collective) while the singular value is overtly coded (unit). Thus, Welsh grammatically recognizes three classes of nouns: singular/plural, collective/unit, and noncountable. I discuss the different classes in turn.

COUNTABLE AND NONCOUNTABLE NOUNS. Grammars of Welsh typically distinguish countable and noncountable nouns based on contrastive morphological coding for singular and plural values. Countable nouns regularly show an alternation between singular and plural values, which correspond to different codings. Noncountable nouns, however, are described as being 'not usually found in the plural' (King 2003:35) or as those nouns that 'do not normally have number contrast' (Jones & Thomas 1977:161). Examples of countable and noncountable nouns are given in 1 and 2, respectively.

(1) Countable (plural-coded) nouns: *afal*/*afalau* 'apple/apples', *cadair*/*cadairiau* 'chair/chairs', *dyn*/*dynion* 'man/men', *merch*/*merched* 'girl/girls', *ton*/*tonnau* 'wave/waves'

(2) Noncountable nouns: *glo* 'coal', *'menyn* 'butter', *mêl* 'honey', *llefrith* 'milk', *llwch* 'dust'

This contrast between countable and noncountable nouns is also supported by cooccurrence behavior with different quantifiers. The examples in 3–4, adapted from Jones & Thomas 1977:175–76, show that certain quantifiers in Welsh indicating cardinality are not acceptable with nouns designating substances.[2]

(3) mwyafrif o 'r llyfr-au/*eira
majority of the book-PL/snow
'majority of the books/*snow'

(4) amryw o 'r llyfr-au/*eira
several of the book-PL/snow
'several of the books/*snow'

So far, the grammatical contrast in English between countable and noncountable nouns finds a parallel in Welsh, further strengthened by a parallel notional contrast, at least for natural concrete entities: countable nouns correspond to entities that are in some sense individuals, while noncountable nouns correspond to material stuff, namely substances and liquids.

COLLECTIVE NOUNS. Countable nouns can be further divided according to the mode of their morphological coding: some nouns possess morphological coding of the plural, while for others the singular value is overtly coded. Grammars of Welsh have not settled on a term for this class, but I follow King's (2003) terminology and designate it as the COLLECTIVE/UNIT distinction and call the morpheme that codes the singular value the SINGULATIVE.

Examples 5–6 (Jones & Thomas 1977:157–58) demonstrate how Welsh nominals may differ in their default form: for *afal* 'apple', the singular is the default (uncoded) form and the plural is coded, while for *adar* 'birds', the plural is the default form and the singular is coded. The number values of these forms can also be observed through pronominal agreement: the pronoun in the responses in 5–6 agrees in number with the noun in the question, both of which are in bold. (The singulative morpheme is glossed as SING as opposed to SG, which glosses the singular.)

[2] The glossing here follows the Leipzig glossing rules. Additional glosses used are: 3DIM: three-dimensional class, DETPL: determinate plural, DIM: diminutive, DISTPL: distributive plural, HUM: human, NHUM: nonhuman, PTCL: particle, REDUPL: reduplicated plural, SCM: specific class marker, and SING: singulative.

(5) a. Lle mae 'r **afal**? — Mae **o** ar y bwrdd.
where is the apple.SG — is it on the table
'Where's the apple? — It's on the table.'
b. Lle mae 'r **afal-au**? — Maen **nhw** ar y bwrdd.
where is the apple-PL — are they on the table
'Where are the apples? — They're on the table.'

(6) a. Lle mae 'r **adar**? — Mae **nhw** ar y wal.
where is the bird.PL — is they on the wall
'Where are the birds? — They're on the wall.'
b. Lle mae 'r **ader-yn**? — Mae **o** ar y wal.
where is the bird-SING — is it on the wall
'Where is the bird? — It's on the wall.'

The grammatical number system of Welsh provides a clear instance of a tripartite number system. As displayed in Table 2, three different categories of opposition in terms of countability are manifested: singular/plural, collective/unit, and noncountable.

COUNTABILITY CATEGORY	SINGULAR	PLURAL	GLOSS
Singular/plural	cadair	cadair-iau	'chair'
Collective/unit	cacyn-en	cacwn	'hornet'
Noncountable	llefrith		'milk'

TABLE 2. Grammatical number categories in Welsh.

Descriptions of Welsh's number system inevitably point to a notional distinction accompanying the formal distinction between the singular/plural and collective/unit categories, namely, the distinction between those types of entities that habitually cooccur in the world and those that do not. According to this view, that Welsh has two types of countable nouns is not some morphological oddity, but a systematic distinction. Stolz (2001:65) analyzes the lexical items that make up the collective class in Welsh and characterizes them as those nouns designating entities that are 'saliently perceived as collectivities rather than "individuals" '. Stolz's findings indicate that the collective class is cohesive in terms of its lexical-semantic categorization. His categorization is partially reproduced in Table 3.

SMALL ANIMALS AND INSECTS	*llygod*/*llygod-en* 'mice/mouse', *berdys*/*berdys-en* 'shrimp/shrimp', *cacwn*/*cacyn-en* 'hornets/hornet', *picwn*/*picwn-en* 'bees/bee', *chwain*/*chwann-en* 'fleas/flea', *morgrug*/*morgrug-yn* 'ants/ant'
MIDDLE-SIZED ANIMALS	*piod*/*piod-en* 'magpies/magpie', *moch*/*moch-yn* 'pigs/pig', *cwning*/*cwning-en* 'rabbits/rabbit', *hwyaid*/*hwyad-en* 'ducks/duck'
VEGETATION/CEREALS/FRUITS/VEGETABLES	*dincod*/*dincod-yn* 'seeds/seed', *ceirch*/*ceirch-en* 'oats/an oat', *chwynn*/*chwynn-yn* 'weeds/weed', *cnau*/*cneu-en* 'nuts/nut', *danadl*/*danhadl-en* 'nettles/nettle', *maip*/*meip-en* 'turnips/turnip', *cennin*/*cenhin-en* 'leeks/leek', *afan*/*afan-en* 'raspberries/raspberry', *gellyg*/*gellyg-en* 'pears/pear', *bresych*/*bresych-en* 'cabbages/a cabbage'
GRANULAR MASS	*tywod*/*tywod-yn* 'sand/grain of sand', *marwor*/*marwor-yn* 'embers/an ember', *llwch*/*llych-yn* 'dust/speck of dust'

TABLE 3. Subtypes of collective nouns in Welsh (adapted from Stolz 2001).

Stolz (2001) argues that both this grammatical category and the notional category underlying it have been historically stable and, furthermore, have gone through expansion via borrowing. Stolz (2001) and Jones and Thomas (1977) observe that borrowing in Welsh appears to be affected by whether the referent of the borrowed term is perceived as a collectivity. Some borrowings from English are given in Table 4.

	WELSH	
ENGLISH TERM	COLLECTIVE	UNIT
bricks	brics	bricks-en
figs	ffigys	ffigys-en
gooseberries	gwsberys	gwsberys-en
fowl	ffowls	ffowl(s)-yn

TABLE 4. Borrowings from English into the collective class in Welsh.

The typical pattern of borrowing is that the morphologically uncoded form is borrowed (see Tiersma 1982); however, the borrowings in Table 4 are remarkable in that the plural, and morphologically complex, English form is borrowed as the base form in Welsh. Welsh speakers have apparently taken the occurrence of multiple entities as the basic situation for these nouns, designated in English in the plural, and thus borrowed the plural form and treat it as uncoded, to which the singulative can then apply.

Altogether, the data from Welsh demonstrate that countability distinctions may be richer than a binary countable/noncountable contrast. The morphosyntactic patterns of the language isolate three classes of nouns: singular/plural, collective/unit, and noncountable.[3] Further, the morphological pattern of collective/singulative seems to reflect a categorization of entities in the world.

IMPLICATIONS FOR COUNTABILITY. One may question whether this additional category of collectives is actually relevant for understanding the distinction between countable and noncountable nouns. This is not obvious, since from the morphosyntactic evidence adduced so far, it could be that the collective is just a flavor of the plural. One line of argumentation would be that these nouns are simply countable nouns, as their agreement patterns would indicate, and therefore this distinction, whatever it ultimately may be, is one that does not pertain to noncountable nouns or aid us in understanding the nature of the countable/noncountable divide. Coming from the perspective of a binary count/noncount contrast, one could question whether these three categories could not be wedged into a binary mold. Although that form of reductionism may seem appealing at first as it keeps to a simple two-way contrast, it runs into far too many difficulties to be maintained.

First, consider the notional distinction between countable and noncountable nouns, where countable nouns correlate with individuals and noncountable nouns correlate with nonindividuals. The core classes of entity types realized as noncountable nouns and countable nouns in Welsh and English are essentially identical, namely liquids/substances and individual entities, respectively. Yet the collective class does not fit comfortably in either. Many members of the collective class would make for extremely dubious individuals, as in the case of the cognates of *dust* and *sand*, for instance, a point reinforced by the fact that these nouns are noncountable in English. This evidence is, however, only suggestive.

A more serious objection to conflating the collective class with simple plural nouns is that it is common for collective nouns to make a three-way distinction between collective, singular, and plural values. An example from Welsh is given in 7, showing that cer-

[3] Two additional noun types, not relevant to the points being made here, can also be established through contrastive singular/plural marking and verbal agreement. First, Welsh distinguishes characteristic 'group' nouns (see Landman 1989, Joosten et al. 2007), nouns designating social bodies or organizations. These nouns manifest a singular/plural distinction; however, the singular form requires the plural form of the verb (Jones & Thomas 1977:161). Welsh also distinguishes PLURALIA TANTUM, that is, nouns that only appear in plural form. This class includes nouns such as *nefoedd* 'heaven' or *trigolion* 'inhabitants'. Unlike group nouns or collective/unit nouns, they do not manifest a singular/plural contrast.

tain collective nouns permit pluralization of the singulative form. This potential to make a three-way distinction is not limited to Welsh but also occurs in the related Celtic language Breton, whose number system is similar to that of Welsh. A Breton example parallel to the Welsh in 7 is shown in 8, taken from Press (2009:445), who glosses the pluralized form as 'individualized' to signal its meaning. The collective form and the plural form have distinct meanings: the pluralized singulative form designates a set of individual pieces in contrast to, for instance, a heap of grain. Treating collectives as plurals faces a serious challenge in accounting for these three-way contrasts.[4]

(7) Welsh (Stolz 2001:70)

a.	b.	c.
grawn	gron-yn	gron-ynn-au
grain	grain-SING	grain-SING-PL
'grain'	'a single grain'	'grains'

(8) Breton (Press 2009:445)

a.	b.	c.
deil	deli-enn	deli-enn-où
leaf	leaf-SING	leaf-SING-PL
'leaves' (collective)	'leaf'	'leaves' (individualized)

In sum, the distributional differences reviewed above show three distinct major classes. The implication for theories of countability, then, is that a binary categorization of nouns into countable and noncountable may be necessary, but it is not sufficient. Both a simple grammatical distinction between nouns compatible with number coding and those not and a notional distinction between individuals and nonindividuals oversimplify the typological space. The data from Welsh's collective class also indicate that nouns such as *leaves* and *ants*, which are not canonical examples of countable or noncountable nouns, provide much interest for understanding countability contrasts.[5] These types of nouns reappear in the other languages that are examined in this section.

2.2. THE COLLECTIVE/SINGULATIVE IN NILO-SAHARAN LANGUAGES. Languages of the Nilo-Saharan family, spoken across a large area in Central and Eastern Africa, show a widespread use of a collective/unit distinction in addition to a countable/noncountable distinction. I primarily exemplify this pattern with examples from Turkana, spoken in Kenya, as it has been well described in terms of grammar and vocabulary (Dimmendaal 1983 and Ohta 1989, respectively).

Turkana displays a countable/noncountable distinction: countable nouns have both a singular and plural form, while the form of noncountable nouns is either exclusively singular or exclusively plural.[6] Number is coded in Turkana by both number suffixes and agreement prefixes, which also code gender. The example in 9 displays the singular and plural form of the word *-kòrì* 'giraffe', where the singular form has the singular masculine agreement prefix *e-*, and the plural form has both the plural masculine/neuter prefix *ŋi-* and the plural suffix *-o*. For noncountable nouns, the agreement prefix shows whether the noun has a singular or plural form. Table 5, adapted from Dimmendaal 1983:211, 234, displays examples of countable and noncountable nouns.

[4] Further notional arguments are given in King 2003:48. Nurmio (2016) provides arguments from morphological patterning against conflating plural and collective values, as well as against treating the collective as simply possessing a null suffix with a plural value, as suggested by Awbery (2009). For instance, the *-en*/*-yn* suffix does not apply to regular plurals, and further, both the singular and the collective base forms, but not the plural, may feed derivation.

[5] See also the wide-ranging discussion in Acquaviva 2008.

[6] The contrast between inherently plural and inherently singular noncountable nouns does not appear to be semantically significant, but is rather due to etymology. See Dimmendaal 2000:230 for discussion.

(9) a. e-kòrì
M.SG-giraffe
'giraffe'

b. ŋi-kori-o`
M.PL-giraffe-PL
'giraffes'

COUNTABILITY CATEGORY	SINGULAR	PLURAL	GLOSS
Countable	a-mosiŋ´	ŋa-mósíŋ-o	'rhinoceros'
	a-rɪɛti´	ŋa-ríét-a	'desert'
Noncountable	a-siŋoní	—	'sand'
	—	ŋa-kipí	'water'

TABLE 5. Countable and noncountable nouns in Turkana.

Countable nouns are further classified by their coding of singular and/or plural interpretations. Table 6 displays the three patterns in Turkana. The singular/plural class codes the plural, while the collective/singulative class codes the singular value via a singulative marker. The third class is known as the 'replacement' pattern (Dimmendaal 1983), where both the singular and plural values are overtly coded. While the exact morphosyntactic and lexical details differ from language to language, Nilo-Saharan languages consistently distinguish these three types of singular/plural patterns. The focus here is on the contrast between the singular/plural class and the collective/singular class, setting aside the replacement pattern since it does not bear on countability per se.

CODING TYPE	SINGULAR	PLURAL	GLOSS
Singular/plural	a-mosiŋ´	ŋa-mósíŋ-o	'rhinoceros'
Collective/singulative	ɛ-sɪkɪn-a´	ŋɪ-sɪkɪn´	'breast'
Replacement	e-kúk-ut	ŋi-kúku-i´	'chicken'

TABLE 6. Types of countable nouns in Turkana.

As in Welsh, there is a clear lexical semantic generalization underlying the collective/singulative class which stands in contrasts to the singular/plural class. Dimmendaal (2000:229) observes that 'entities congregated in large numbers or quantities tend to be morphologically unmarked in the plural in Nilo-Saharan languages', and for these nouns the singulative codes 'an individuated item from a collective or group'.

Despite their geographical and genetic separation, Nilo-Saharan and Celtic number systems show substantial overlap in the types of entities inhabiting their collective/singulative classes. Table 7 shows examples from the lexicon of Turkana, which include nouns of the same entity types found in Welsh, along with two further entity types—types of people and paired/grouped body parts.

TYPES OF PEOPLE	*e-sebey-it*/*ngi-sebey* 'Sebey person/Sebey people', *e-turkana-it*/*ngi-turkana* 'Turkana person/Turkana people'
PAIRED/GROUPED BODY PARTS	*e-pon-ol*/*ngi-pon* 'lip/lips', *e-kyal-ai*/*ngi-kyal* 'tooth/teeth', *a-ki-t*/*nga-ki* 'ear/ears', *e-gec-ot*/*ngigece* 'wrist/wrists'
SMALL ANIMALS AND INSECTS	*e-suro-t*/*ngi-suro* 'mosquito/mosquitos', *e-kur-ut*/*ngi-kur* 'maggot/maggots', *e-lac-it*/*ngi-lac* 'bed louse/bed lice'
MIDDLE-SIZED ANIMALS	*a-taruk-ot*/*nga-taruk* 'vulture/vultures', *a-toow-at*/*nga-toowa* 'duck/ducks'
VEGETATION/CEREALS/FRUITS	*a-tur-ot*/*nga-tur* 'flower/flowers', *e-mar-et*/*ni-mare* 'bean/beans', *e-nga-it*/*ngi-nga* 'blade of grass/grass', *a-kuy-en*/*nga-kuui* 'leaf/leaves'
GRANULAR MASS	*e-cok-et*/*ngi-cok* 'seed/seeds', *e-tab-a*/*nga-tab* 'piece/pieces of tobacco'

TABLE 7. Subtypes of collective nouns in Turkana (extracted from Ohta 1989).

In sum, the Turkana grammatical number system, and more broadly those of Nilo-Saharan languages, makes a three-way distinction among entity types that are not countable, those that typically come as multiple and related entities, and those that typ-

ically come as individual entities. This three-way distinction can be extrapolated from the grammatically recognized categories of countability, namely singular/plural, collective/singulative, and noncountable. These languages then demonstrate again that the countable and noncountable contrast must only be a partial description of the countability distinctions languages may express.

2.3. THE COLLECTIVE/SINGULATIVE IN MALTESE. Maltese, a Semitic language historically in intensive contact with European languages, also displays a tripartite number system, but differs in interesting ways both in terms of behavior with cardinal modifiers and in terms of the lexical semantic domains involved. Maltese has a countable/noncountable contrast that is witnessed by the usual core distributional properties. Countable nouns code a plural interpretation and allow modification by quantifiers implicating plurality and cardinal determiners. Cardinal modification from 'two' to 'ten' requires the plural form of the noun, while cardinal modification from 'eleven' onward requires the singular form. Table 8 shows the paradigm for countable nouns, adapted from Mifsud 1996. In contrast, noncountable nouns typically have only a single form, as in *arju* 'air', and do not permit modification by quantifiers implicating plurality or by cardinal determiners.

SINGULAR	PLURAL	INDEFINITE QUANTIFIERS	CARDINALS 2–10
ˈskola	ˈskeyyel	ˈhafna ˈskeyyel	ˈerba˚ ˈskeyyel
'school.SG'	'school.PL'	'many school.PL'	'four school.PL'
ˈrağel	irˈğil	ˈhafna irˈğil	ˈerba˚ irˈğil
'man.SG'	'man.PL'	'many man.PL'	'four man.PL'

TABLE 8. Countable noun paradigm in Maltese.

As in Welsh and Turkana, Maltese possesses a collective/singulative contrast; however, unlike in Welsh and Turkana where the collective is formally a plural, for instance, in terms of agreement, the Maltese collective is zero-coded in the collective interpretation and formally singular and masculine. The unit interpretation is achieved through the singulative suffix *-a*, which also renders the noun feminine yet still singular.

The combinatoric patterns with quantifiers add an additional layer of complexity to the collective's nominal paradigm. In contrast to countable nouns, which have just one plural form, collective nouns have an additional plural form when used with lower cardinal modifiers (2–10), known as the DETERMINATE PLURAL. For quantifiers of indefinite quantity, the collective form is used. The paradigm is given in Table 9, adapted from Mifsud 1996.

COLLECTIVE	UNIT	INDEFINITE QUANTIFIERS	CARDINALS 2–10
hût	ˈhûta	ˈhafna ˈhût	ˈerba˚ hûˈtit
'fish'	'fish.SING'	'many fish'	'four fish.DETPL'
ˈnemel	ˈnemla	ˈhafna ˈnemel	ˈerba˚ nemˈlit
'ant'	'ant.SING'	'many ant'	'four ant.DETPL'

TABLE 9. Collective noun paradigm in Maltese.

The types of entities that exhibit the collective/singulative paradigm in Maltese are given in Table 10. There is substantial overlap with the types of entities seen in Welsh and Turkana—the core members of vegetation, insects, and granular aggregate are present. Yet there are some striking differences—the category of types of people found in Turkana (as well as in Arabic) is absent, along with middle-sized animals that typically come in groups, as seen for Welsh and Turkana. If higher-level animates are absent in nominal domains of the collective in Maltese, the categories of foodstuffs (*bread*, *cheese*) and materials (*iron*, *wood*), in contrast, are novel to the discussion so far.

PAIRED ENTITIES	*zarˈbûn* ‘shoes’, *karkur* ‘slippers’, *taraǧ* ‘stairs’, *bib* ‘doors’
SMALL ANIMALS AND INSECTS	*hut* ‘fish’, *tayr* ‘fowls’, *nemel* ‘ant’, *dubˈb-in* ‘flies’, *nahal* ‘bees’
VEGETATION/CEREALS/FRUITS	*ʾahm* ‘corn’, *eneb* ‘grapes’, *tin* ‘figs’, *lumi* ‘lemon’
GRANULAR AGGREGATES	*ˈramel* ‘sand’, *trab* ‘dust’
FOODSTUFFS	*hobz* ‘bread’, *ˈǧobon* ‘cheese’, *cokkor* ‘sugar’
MATERIALS	*haˈdid* ‘iron’, *spag* ‘string’, *inˈy-am* ‘wood’

TABLE 10. Subtypes of collective nouns in Maltese (after Mifsud 1996).

The literature on Maltese has often noted that the collective/singulative class spans entity types that are both countable nouns (*shoes*) and logically noncountable nouns (*iron*, *wood*) (Sutcliffe 1936, Borg 1980, Mifsud 1996:32). A closer look at the data shows that the singulative has two distinct uses, which depend on the type of entity. When the singulative combines with an aggregate noun, the singulative designates one member of the aggregate. When the singulative combines with foodstuff nouns, as well as for certain material nouns, the singulative designates one conventional portion, where, for example, the unit interpretation of ‘bread’ results in ‘a loaf’. These different interpretations are shown in Table 11. The countable interpretations of the entities in the foodstuff and material categories are reminiscent of interpretations attributed to the UNIVERSAL PACKAGER in English (see Pelletier 1975), such as when *a water* designates a glass or bottle of water. Like the universal packager, the use of the singulative with noncountable nouns is restricted just to those that have a conventionalized portion interpretation, and it may not cooccur with any arbitrary noncountable noun.[7]

COLLECTIVE	GLOSS	UNIT	GLOSS
baqar	‘cattle’	baqra	‘a cow’
nahal	‘bees’	nahla	‘a bee’
sigar	‘trees’	sigra	‘a tree’
taraǧ	‘(a flight of) stairs’	tagra	‘a step’
zarˈbûn	‘(a pair of) shoes’	zarˈbûna	‘a shoe’

MATERIAL	GLOSS	CONVENTIONAL PORTION	GLOSS
hobz	‘bread’	ˈhobza	‘a loaf of bread’
ˈǧobon	‘cheese’	ˈǧobna	‘a cheese’
sapûn	‘soap’	sapûna	‘a bar of soap’

TABLE 11. Two interpretations of the singulative in Maltese: member or conventional portion.

The grammatical number system in Maltese, as in Welsh, shows every sign of being productive. Evidence from borrowing shows that entities of the appropriate type are simply integrated into the collective category. Mifsud (1996:44–45) discusses borrowings such as the Maltese term for ‘sandals’, based on the Italian plural *sandali*, which is integrated into the collective/singulative paradigm as *sandli*/*sandl-a*. Similarly, the term for ‘brick’, based on English *bricks*, appears as *briks*/*briks-a*. As in Welsh, here it is the plural form in the source language that is borrowed.

Taken together, the data from Maltese demonstrate yet another instance where a three-way contrast is in effect between types of entities that are not countable, those that

[7] Since *soap* and other material nouns take singulative/plural coding in Maltese, it is tempting to analyze these nouns as countable rather than noncountable. Yet other diagnostics demonstrate clearly that the class of nouns taking the singulative is formed from two types of members, namely logically countable entities and logically noncountable entities. First, adjectival modifiers that presuppose individuals, such as *big* or *square*, combine with collectives composed of naturally countable entities, such as *hut* ‘fish’, but not those composed of material, such as *kartun* ‘cardboard’, unless in combination with the singulative or determinate plural (Gil 1996). See Grimm 2012c for further examples and discussion.

typically come in groups, and those that are typically individual objects. Further, the lexical semantic domains of each category overlap substantially with those found in Welsh and Turkana.

2.4. DAGAARE: BEYOND TRIPARTITE NUMBER SYSTEMS. Dagaare, a Gur language spoken in Northern Ghana, presents an even more complicated grammatical number system.[8] Dagaare, as spoken in the Central Dialect in and around the town of Jirapa, Ghana, not only grammatically recognizes countable and noncountable nouns, but also makes finer distinctions within both classes.

COUNTABLE/NONCOUNTABLE CONTRAST. The countable/noncountable contrast in Dagaare can be established by using some of the core distributional properties. First, countable nouns display a singular/plural contrast, while noncountable nouns do not have a plural form, as shown in Table 12. Similarly, countable, but not noncountable, nouns combine with cardinal modifiers, as shown in 10.[9]

NOUN	PLURAL FORM	GLOSS
tìɛ́	tììrí	'tree'
pìɛ́	pɛ̀rí	'basket'
dʊ̀ɔ́	dòrí	'pig'
kʊ̀ɔ́	—	'water'
kãã̀	—	'oil'
zĩĩ̀	—	'blood'
sáálá	—	'charcoal, coal'

TABLE 12. Nouns with and without a singular/plural contrast in Dagaare.

(10) dò-rí à-yí
pig-PL NHUM.PL-two
'two pigs'

The use of cardinal modification with noncountable nouns is not allowed even with shifts of meaning of the sort associated with packaging found in English, such as *two glasses of water*.[10] More broadly, contextual variability familiar from the discussion of English in the form of grinding, packaging, and sorting contexts was not available in Dagaare. For instance, Dagaare speakers strongly reject the use of nouns such as *báá* 'dog' in grinding contexts.[11]

Although noncountable nouns do not have a plural form in general, they may combine with other number suffixes. First, both countable and noncountable nouns may optionally combine with a distributive plural suffix *-ree*. Examples of nouns with the distributive plural are given in Table 13. The distributive plural, however, is not able to

[8] The data in this section were gathered over the course of seven months of fieldwork conducted by the author in 2008 and 2011. For more general information on Dagaare grammar, see Bodomo 1997, 2000.

[9] Cardinal modifiers above 'two' appear with agreement prefixes, which agree in number and distinguish human and nonhuman referents. The prefix *a-* agrees with nonhuman plural nouns.

[10] Only one noncountable noun, *dãã̀*, referring to the local alcoholic beverage, was accepted by speakers in combination with cardinal modifiers under a packaging reading. All other nouns were strongly rejected.

[11] Sentences constructed to elicit grinding interpretations, such as in (ia), were judged unacceptable by Dagaare speakers, who instead required modification, as shown in (ib).

(i) a. ??baa paale la a sori (zaa)
dog full PTCL DEF road (all)
'There was dog all over the road.'
b. ba-nɛnɪ paale la a sori (zaa)
dog-meat full PTCL DEF road (all)
'There was dog meat all over the road.'

combine with cardinal numbers, as shown in 11. Noncountable nouns then, while accepting a type of indefinite plurality, are not equivalent to canonically countable nouns that do permit cardinal modification, as was shown in 10.

NOUN	DIST. PL.	GLOSS
kʊ̀ʊ́	kɔ̀nnéɛ́	'water; types of waters; water in different locations'
mʊ́ɔ	mʊ́ɔnéɛ́	'grass; types of grasses; grass in different locations'

TABLE 13. Distributive plural in Dagaare.

(11) *kɔ́nnéɛ́ à-yí
water.DISTPL NHUM.PL-two
'two waters'

VARIETIES OF NONCOUNTABLE NOUNS. Noncountable nouns can be further classified as to whether they permit a singulative suffix *-ruu*, designating 'a piece of'. Unlike the singulative morphemes discussed so far, the distribution of the singulative in Dagaare is very limited: it combines with fewer than 100 nouns, designating primarily granular aggregate terms or foodstuffs (*sérímáání* 'pepper', *kpéɛ́* 'malt', *múó* 'grass', *kàmáání* 'corn').[12]

NOUN	SINGULATIVE	DIST. PL.	GLOSS
kʊ̀ʊ́	—	kɔ̀nnéɛ́	'water/—/(types of) waters'
mʊ́ɔ	mʊ́ɔrʊ́ʊ	mʊ́ɔnéɛ́	'grass/blade of grass/grasses'

TABLE 14. Number-marking possibilities for noncountable nouns in Dagaare.

The distribution of *-ruu*, shown in Table 14, implies that the noncount nouns in Dagaare can be divided into those that accept *-ruu* and those that do not, a morphosyntactic division that corresponds to a difference in entity type: granular aggregates as opposed to liquids and substances.

VARIETIES OF COUNTABLE NOUNS. Count nouns in Dagaare have singular and plural forms, yet which is morphologically coded varies by noun, giving rise to a pattern known as 'inverse number marking'.[13] Table 15 contrasts the nouns for 'child' and 'seed': they share the same stem, yet for 'child' the morpheme *-ri* codes the plural interpretation, while for 'seed' it codes the singular interpretation. Additional examples in Tables 16 and 17 demonstrate *-ri* coding plural and singular interpretations, respectively. Note that *-ri* undergoes assimilation following nasals and liquids, resulting in the allomorphs *-nI* and *-lI*, respectively. (For vowels, I use capital letters as a variable for ±ATR values—that is, *-rI* ranges over *-rɪ* and *-ri*.)

GLOSS	SINGULAR	PLURAL	STEM
'child'	bíé	bíírí	bì-
'seed'	bírì	bíè	bí-

TABLE 15. The inverse number-marking paradigm in Dagaare.

-V SINGULAR	*-rI/-nI* PLURAL	GLOSS
tìɛ́	tììrí	'tree'
gbìé	gbèrí	'forehead'
pìɛ́	pɛ̀rí	'basket'
nàŋá	nànní	'scorpion'

TABLE 16. Marked plural nouns in Dagaare.

[12] There also appears to be a use of *-ruu* among some speakers to indicate a packaging reading, where for instance *kpéérʊ́ʊ́* designates one sack of malt rather than a piece.

[13] See Grimm 2012b for further discussion of Dagaare's inverse number-marking system and its implications for the understanding of plurality and markedness.

-rI/-nI SINGULAR	-V PLURAL	GLOSS
lúgrí	lúgó	'prop, pillar'
nyágrí	nyágá	'root'
fìlí	fìlé	'sore'
íílí	íílè	'horn'

TABLE 17. Marked singular nouns in Dagaare.

Agreement with demonstratives and number words establishes that the singular and plural forms code bona fide singular and plural interpretations. The demonstrative agreement patterns for plural-coded and singular-coded nouns are shown in 12 and 13, respectively: the demonstrative has a singular (*ŋâ*) and plural (*-má*) form, the plural further distinguishing via a prefix between human (*ba-*) and nonhuman (*a-*) referents.

(12) a. bíé ŋâ
child.SG DEM.PROX.SG
'this child'

b. bíí-rí bà-má
child-PL HUM.PL-DEM.PROX.PL
'these children'

(13) a. bí-rì ŋâ
seed-SG DEM.PROX.SG
'this seed'

b. bíè à-má
seed.PL NHUM.PL-DEM.PROX.PL
'these seeds'

Number words designating 'two' or more select for plural nouns and again take plural agreement prefixes. The word designating 'one', *yénì*, does not show agreement but acts as a modifier of the noun, directly attached to the stem. The number-word agreement patterns for plural-coded and singular-coded nouns are shown in 14 and 15, respectively.

(14) a. bì-yénì (bì- + yénì)
child-one
'one child'

b. bíí-rí bá-yì
child-PL HUM.PL-two
'two children'

(15) a. bí-yénì (bí- + yénì)
seed-one
'one seed'

b. bíè à-yí
seed.PL NHUM.PL-two
'two seeds'

As one might suspect in light of the examinations of the Welsh, Turkana, and Maltese number systems, a systematic examination of the lexicon developed in the field shows reliable asymmetries across different nominal domains. Figure 1 shows counts in the lexicon for the domains of animals, insects, trees, vegetation, and tools. The dark-shaded bars indicate nouns with the singular unmarked, while the light-shaded indicate those with the plural unmarked; for example, for mammal, forty-three entries are unmarked in the singular and five in the plural.[14]

Overall, nouns for higher-level animates—namely mammals, birds, and reptiles—are typically unmarked in the singular; however, the majority of nouns for insects have a plural that is unmarked. Similarly, nouns for trees are typically unmarked in the singular, while most nouns for vegetation are unmarked in the plural. By way of comparison, the graph includes data for nouns designating tools, which designate items that could reasonably be hypothesized as entities that naturally occur as self-standing individuals. These nouns showed a strong tendency toward being unmarked in the singular.

An asymmetry in the domain of body parts is displayed in Figure 2. Nouns designating body parts that inherently come in pairs or groups, such as 'eye' or 'rib', are more likely to be unmarked in the plural; yet nouns designating body parts that are inherently singular, such as 'head', are more likely to be unmarked in the singular.

[14] In these counts, I exclude derived forms, since they follow their own patterns, which tends to obscure any generalization.

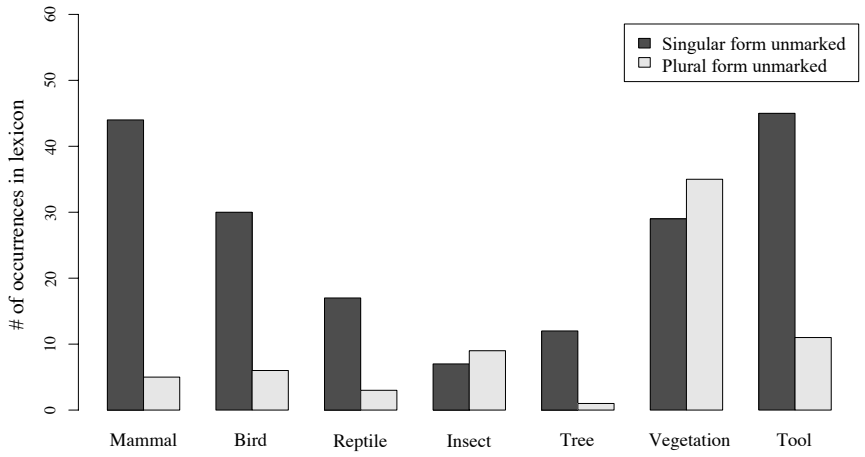

FIGURE 1. Number marking across semantic domains in Dagaare.

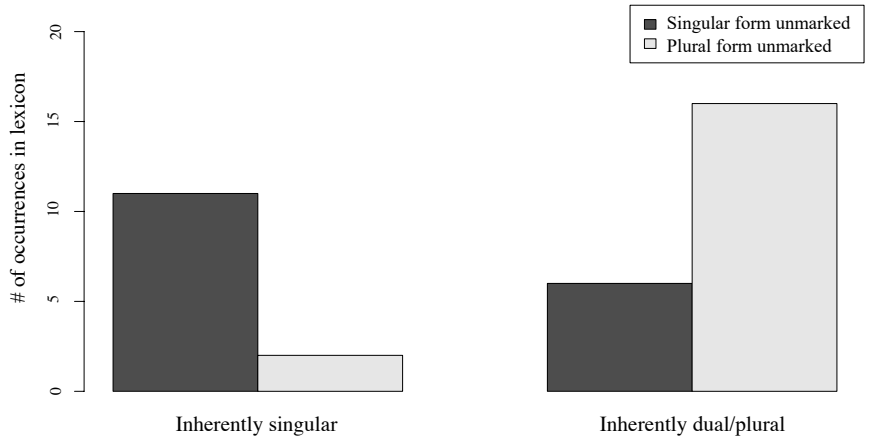

FIGURE 2. Number marking and inherent plurality in the domain of body parts in Dagaare.

Thus, Dagaare makes a distinction similar to what has been shown for the other collective/singulative systems: nouns that are likely to cooccur or form a natural pair or group, shown in Table 18, are distinguished from nouns that are likely to be seen as individual entities. The manner of coding this distinction—through inverse number marking—is the novel aspect of the Dagaare system.

PAIRED BODY PARTS	*dumi/duno* 'knee/knees', *gberı/gbee* 'leg/legs', *sıngkpıgrı/sıngkpıgɛ* 'hip/hips', *nukumi/nukuno* 'fist/fists', *yagrı/yaga* 'cheek/cheeks'
INSECTS/REPTILES/BUGS	*dunduli/dundulo* 'maggot/maggots', *soori/sobo* 'locust/locusts', *yɔɔnı/yɔɔmɔ* '(a type of small) frog/frogs'
VEGETATION/CEREALS/FRUITS	*kamaanı/kamaana* 'corn', *lungguri/lungue* 'lemon/lemons', *guori/goe* 'cola nut/cola nuts'

TABLE 18. Subtypes of unmarked plural nouns in Dagaare.

Altogether, the distribution of the different morphological markers implicates five nominal types that receive grammatical recognition in Dagaare: substance, granular aggregate, collective aggregate, individual, and human. Once again, a purely countable/noncountable distinction would not capture all of the nuances that Dagaare manages to express.

2.5. SUMMARY: COMPARING GRAMMATICAL CLASS, CODING, AND ENTITY TYPE. The data presented so far in this section would appear to add arguments, alongside those discussed in §1, against the hypothesis that semantic classification underlies the count/mass distinction. If English and French cognates (such as *hair* and *cheveux*) do not fall cleanly into the same countability categories, expanding the crosslinguistic database only exacerbates the problem. These mismatches, however, do not prove that semantic classification is completely absent—they simply indicate that there is no one-to-one correspondence between entities and their countability classification valid across all languages. Yet from what we know about other grammatical categories, such as grammatical relations or definiteness, it would be extremely surprising if there was such agreement among languages—total alignment of the content of grammatical categories across languages is rare.

Despite these mismatches, much overlap is present in these different grammatical number systems. The goal here is to illuminate where languages do converge and provide a framework that represents this information in a predictive manner. This clearly cannot be achieved at the level of particular entities in the world, as cognates with opposite countability classification demonstrate, and I argue in this section that this does not even hold at the level of types of entities. This in turn motivates couching the explanation of countability in more abstract properties related to individuation, which occupies §3.

The different languages examined show substantial variation in the inventory of entity types appearing in their collective/singulative classes. A first question to ask is if any systematicity can be found among this variation. Table 19 plots the relation between entity types and their manner of coding in languages with a singulative marker. The entity types given here have no formal status but are simply descriptive categories based on the discussion in this section. The rows of the table give the entity types, while the columns give the languages. Each cell specifies the coded value, such as singular by SG; additionally, shaded cells signify entity types falling in collective/unit classes in the different languages. (Dagaare codes the singular value for two categories of nouns: those with the singular marked by *-ri* and those with the optional singulative *-ruu*, which are shaded in dark and light gray, respectively.)

ENTITY TYPE	CODING			
	WELSH	TURKANA	MALTESE	DAGAARE
TYPES OF PEOPLE	pl	sg	pl	pl
MIDDLE-SIZED ANIMALS	sg	sg	pl	pl
PAIR/GROUPED BODY PARTS	pl	sg	pl	sg
SMALL ANIMALS AND INSECTS	sg	sg	sg	sg
VEGETATION/CEREALS/FRUITS	sg	sg	sg	sg
GRANULAR AGGREGATES	sg	sg	sg	0/sg
FOODSTUFFS	0	0	sg	0/sg
LIQUIDS	0	0	0	0

TABLE 19. Entity types and their number coding in Welsh, Turkana, Maltese, and Dagaare.

The relation between entity types and their coding in Table 19 reveals that some entity types are coded via a collective category more often than others across the different languages. While insects and vegetation are uniformly treated as members of the collective category, higher animates are less often treated as such. Granular aggregates are always distinguished from liquids, although foodstuffs are coded variably. This table indicates that there appears to be a core set of entity types, namely insects and vegetation,

where if a language grammatically recognizes a collective category, entities from these types will fall under it.

The information in Table 19 further suggests an ordering of the entity types based on their coding preferences across these different languages. Ranking the manners of coding, from noncountable to countable nouns with collective nouns in the middle, namely ZERO CODING (0) < SINGULAR CODING (SG) < PLURAL CODING (PL), and then counting the number of cells in a row associated with each coding establishes an ordering, the scale in 16.[15] This scale is not strictly ordered, since some entity types are ranked equally (as indicated by '≤').

(16) liquids < foodstuffs < granular aggregates < vegetation/cereals/fruits ≤ insects < small animals < paired/grouped body parts ≤ middle-sized animals < types of people < individuals

The ordering in 16 appears to reflect the accessibility of the unit interpretation. The higher in the scale an entity type is, the more salient the unit interpretation. The entity types lowest on the scale have zero-coded nouns, where units are either not coded directly or not uniformly coded. For instance, in Welsh, *llefrith* 'milk' does not directly code a unit, and a unit interpretation can only be brought about by using additional material, such as a measure phrase. In contrast, *cacwn* 'hornet', a member of the insects entity type higher up the scale, has morphological means for regularly expressing a unit, namely the singulative. In Dagaare, for some zero-coded nouns, namely granular aggregates and foodstuffs, a unit interpretation can be specified through the singulative morpheme, but this specification is not obligatory or regular. For singular-coded nouns, the expression of the unit value is regularly made available, but it is not the default. For plural-coded nouns, the expression of the unit value is the default.

As compelling as the scale in 16 might be, it falls short in several ways. First, that some members of an entity type belong to, for instance, the collective/singulative class does not indicate that they all do. As the lexicon counts for Dagaare in §2.4 showed, these lexicalization patterns aligning entity type and grammatical coding are strong correlations, but not absolutes. Further, the scale in 16 gives the impression that the size of an entity is a factor relevant to its countability classification, yet, as is discussed below, the influence of size is epiphenomenal. Finally, one can question how well this scale will generalize beyond the four languages upon which it is constructed. From other data related to singulative marking discussed in §4 and from the additional literature on collectives/singulatives referenced there, the broad trend of the scale in 16 would appear to generalize—yet only a larger typological study could confirm how well.

Despite these shortcomings, Table 19 and its associated scale indicate that grammatical number appears to be responsive in some manner (i) to ontological distinctions among different entity types, such as substances versus entities with discrete, and potentially countable, elements and (ii) to the accessibility of units. In sum, while a noun's countability classification cannot be directly deduced from the type of entity, grammatical number does indeed seem to pattern with what nouns designate.

The next section ties accessibility of the unit interpretation to the more abstract notion of individuation. As discussed in §1, the hypothesis that countability marking reflects something about the world is one of the central points of controversy in the countability literature. Increasing evidence from psycholinguistic studies argues for a

[15] To exemplify, the ordering among the following entity types—liquids < foodstuffs < granular aggregates < middle-sized animals < types of people—reflects the values in the cells of their rows, namely {0, 0, 0, 0} < {0, 0, sg, sg} < {sg, sg, sg, sg} < {sg, sg, pl, pl} < {sg, pl, pl, pl}.

relation between speakers' use of grammatical number and how they interact with and view the world. I now turn to discussing some evidence that both supports such a relation, and yet indicates that it is not always a simple relation.

3. THE SCALE OF INDIVIDUATION. The ordering on the scale in the preceding section appears to be linked to the propensity for the entity described by the noun to occur as an individual—a property known as individuation, which has attracted equal parts interest and scrutiny in the psychological and philosophical literatures. Most broadly, the thesis of individuation relates cognitive or perceptual qualities of entities to the grammatical realization of count and noncount nouns. The simplest hypothesis posits a direct correspondence between count nouns and reference to 'individuals' in the world and noncount nouns and reference to 'nonindividuals'. On a strong version of this correspondence theory, language users should 'conceptualize the referents of count nouns as distinct, countable, individuated things and those of mass nouns as nondistinct, uncountable, unindividuated things' (Wisniewski et al. 1996:271). This dichotomous view of individuation is discussed widely in the psycholinguistic literature (see, among others, Bloom 1990, 1994, Wisniewski et al. 1996, Barner & Snedeker 2005). I argue that, given the data from §2, a dichotomous distinction is too coarse and that a scalar approach to individuation is necessary.

This section first discusses evidence from the psycholinguistic literature that (i) grammatical countability is related to the object/substance distinction, and thereby an individual/nonindividual distinction, and (ii) that individuation is related to factors beyond only the object/substance distinction. Considering these factors together implies a scalar approach to individuation.

3.1. INDIVIDUATION AND THE OBJECT/SUBSTANCE DISTINCTION. Researchers in language acquisition have brought together a body of evidence that children are sensitive to an object/substance distinction prior to learning language. Soja and colleagues (1991) conducted a 'word extension task', which demonstrates that children distinguish between solid objects and shapeless substances even before they have acquired a syntactic mass/count distinction. If infants learn a name for a novel solid object, they extend that name to another object of the same shape, but not to one of the same material. If they learn a name for nonsolid material, they extend the name to an item formed from the same material, not to one with a similar shape. For further results indicating the robustness with which children distinguish between solid objects and shapeless substances, see Landau et al. 1988, Imai & Gentner 1997, Samuelson & Smith 1999, and references therein.

Prasada and colleagues (2002) provide evidence that multiple factors influence the countability classification of nouns beyond the object/substance distinction. Adult speakers confronted with novel entities were more likely to use countable nouns to describe them if the entities possessed a regular rather than irregular shape. They further observed that if speakers saw a set of entities with the same irregular shape, then they were more likely to use a countable noun to describe the entity than if they saw just a single irregular-shaped entity. Additionally, if entities were shown to be designed for some function, then they were also more likely to be described by a countable noun. Since all of the stimuli were discrete objects, it is apparent that countability categorization involves more than a simple discrete versus nondiscrete distinction. Instead, Prasada and colleagues (2002) provide evidence that speakers categorize entities as (non)countable according to whether they CONSTRUE an entity, that is, some portion of the world, as an individual worth counting on the basis of multiple factors.

Further studies demonstrate that multiple factors bear on countability classification. Middleton and colleagues (2004) examined two hypotheses from the work of Wierzbicka (1988). The first hypothesis, DISTINGUISHABILITY, proposes that nouns designating entities for which the constituents are more easily distinguishable are more likely to be used as countable nouns, while those nouns designating entities for which the constituents are not easily distinguishable will be used as noncountable nouns. Middleton and colleagues (2004) presented subjects with two displays of novel aggregates, each containing forty elements, which varied in the spatial proximity among the elements (CLOSE versus APART) and the size of elements (LARGE versus SMALL). For example, a subject would see two sets of an element: for one set, each element was large and spatially separated from the others; for the other set, each element was small and spatially contiguous with the other elements. Participants then matched a nonce countable phrase (*These are worgels*) or noncountable phrase (*This is worgel*) with one of the displays. Subjects' choices of countable or noncountable phrase were significantly influenced by the spatial proximity, but not by the size, of the elements.

Middleton and colleagues (2004) also investigated whether the canonical MANNER OF INTERACTION with an entity influences countability. Wierzbicka argues that entities with which one interacts one by one will be described by countable nouns, while entities with which one interacts in quantities, rather than individually, will be described by noncountable nouns. Middleton and colleagues (2004) presented subjects with a novel aggregate, yellow decorative coarse-grained sugar, in a cardboard box. As hypothesized, when subjects had to interact with individual grains, under the guise of engaging in a game, they were significantly more likely to describe the novel aggregate with a countable phrase rather than a noncountable phrase.

Real-world cases provide supporting evidence for the role of interaction. Zwicky (2001) discusses nouns that, while initially categorized as noncountable, are recategorized as countable nouns due to changes in circumstances of their use. For instance, the word *chad*, which designates the tiny pieces of paper left over from punch-card ballots, underwent a shift in countability status during the tight 2000 US presidential election. It became critical after the election to count each vote in certain areas by determining whether the ballot had been completely punched through, and accordingly to examine the pieces of paper designated by *chad*. Previous to the election, most of those working with punch-card ballots, and punch cards in general, typically used the word as a noncountable noun. During the election, as individual votes were scrutinized, *chad* was also used as a countable noun. This is not surprising if interaction with entities is related to countability: in earlier circumstances, *chad* was equivalent to leftover paper from punch cards, but in the election, each chad became significant, representing a vote. Thus, the factors discussed in this section are not restricted to experimental settings, but provide plausible routes of explanation for shifts in meaning, routes of explanation that are not available for a strictly grammatical account, for instance.

The various experimental results imply a complex relation between entities in the world and their countability classification mediated by individuation factors. Countability classification cannot reduce to only objective ontological properties of entities in the world: many of the effects observed concern not properties intrinsic to entities, but relations those entities are perceived to stand in, such as spatial proximity among similar entities or relations to functions or to common interactions. Accordingly, individuation is related not only to intrinsic features of objects, such as shape, but also to whether elements are recognized as sufficiently independent from one another, either spatially or functionally, to be construed as individuals. Yet countability classification cannot be

completely independent of the external world, for how things are construed depends on features of the external world and our interaction with them. The various pieces of experimental evidence indicate that when features of an entity in the world are changed, this may affect its likelihood of being classified as countable or not.

As we move on to consider how individuation factors relate to nominal meaning, it is important to recognize a difference between the experimental contexts in this section and how individuation is encoded in lexicalized meaning. Within the experiments in, for instance, Prasada et al. 2002, participants must reason from their observations about a novel object to a countability classification. I do not assume that such a process underlies daily language use. Rather a noun, an entity description, encodes a construal of the entity it describes. This construal is not calculated on-line, but is rather a generalization over speakers' experiences with entities in the world, which in turn have provided the very properties that lead to the construal of those entities as individuated or not. As the example of *chad* indicates, this relation between the entity and how it has been encoded is not static, but is open for renegotiation when speakers' experiences change sufficiently.[16]

3.2. A SCALAR VIEW OF INDIVIDUATION. I now argue that these individuation factors relate to countability classification in a structured way. Given the various distinct factors present in the experimental evidence just reviewed, a view on individuation that divides entities into individuated and nonindividuated would appear to be an oversimplification. Also, from the perspective of countability classes, which is the grammatical classification that individuation is invoked to explain, a simple contrast between individuated and nonindividuated will be too coarse grained. Even for English, often considered to have a binary countable/noncountable distinction, the differential behavior of nouns across a variety of diagnostics demonstrates that there are 'degrees of countability' that are grammatically detectable (Allan 1980).[17] Thus, from different perspectives, it is plausible that individuation is better viewed as a scalar phenomenon rather than one that can be reduced to a simple binary feature ±INDIVIDUATED.

I now bring together and reformulate the factors discussed in §3.1. These factors do not form an exhaustive inventory of the relevant factors, and I consider a further factor of animacy in §4.2 below, yet they are sufficient to arrive at the minimal set of distinctions needed. At this point, I am also still restricting the investigation to natural concrete entities and their canonical interpretations.

I consider the following four individuation factors to guide how nouns are lexicalized with respect to countability: CONSISTENCY OF SHAPE, DEGREE OF INTERACTION, SPATIAL SEPARATION, and VISIBLE PERCEPTIBILITY. All of these factors are considered relative to minimal units. That is, the factor of visible perceptibility expresses the likelihood of visibly perceiving minimal units, high for *dog*, less so for *sand*, and impossible for *oil*. Note that here I have decomposed the factor termed *distinguishability* in Middleton et al. 2004 into two factors, spatial separation and visible perceptibility, for they have distinct effects, despite being correlated.

Interdependencies hold among the factors. A sufficient degree of visual perceptibility is necessary for having a high value for consistency of shape, spatial separation, or de-

[16] Thus, the position here, or invoking individuation generally, should not be equated with an 'internalist' cognitive account with no connection to the world.

[17] Although Allan (1980) was the first to recognize different degrees of countability, this was strictly limited to a characterization of the syntactic distribution of nouns in English. His approach purposefully refrained from making any connection to the semantics of countability (p. 547, n. 3) and thus differs fundamentally from the approach developed here.

gree of interaction. For example, it cannot be determined if an entity has minimal units of consistent shape unless the units are perceptible in the first place. Likewise, if instances of an entity are not spatially separated or, in other words, they are overlapping, then they will also not be visibly perceptible, nor will they have a consistent shape. Additionally, there are cases in which one factor may be more highly weighted than another, as when a high level of interaction with individual units leads speakers to classify entities as countable despite a high level of spatial contiguity, as in the case of the yellow sugar aggregate in Middleton and colleagues' (2004) study or of *chad* in the 2000 US presidential election.

The degree of individuation of a nominal description can be calculated in terms of its composite ranking derived from how it ranks on the different individuation factors. Since each of the factors is continuous, of course infinitely many classes could be formed. Yet certain contrasts are more salient than others, and I establish four classes that account for the primary distinctions encountered: LIQUIDS/SUBSTANCES, GRANULAR AGGREGATES, COLLECTIVE AGGREGATES, and INDIVIDUALS.

An initial contrast is the presence of visibly perceptible minimal units.[18] Least individuated are nouns describing entities without any perceptible minimal units—the standard description in the literature for substances. If an entity does not have perceptible minimal units, the rest of the individuation factors have no traction. Nouns describing entities with perceptible units, but which are low-ranking on the other factors such as more consistent shape, but not typically separated from one another and not typically involved in interactions with individual units, as is the case for granular aggregates such as *rice* or *sand*, are more individuated than substances, but still overall low in individuation. Next, nouns describing entities with perceptible units and mid-level rankings for the other factors, such as being separated from one another but still connected in some fashion, whether spatially near or functionally united, as for collective aggregates such as *ants* or *cherries*, are then again more individuated. Nouns describing entities whose elements are independent from one another, not connected in a regular manner to other elements of the same class, are the most individuated.

I refer to these different classes as INDIVIDUATION TYPES. More precisely, individuation types are equivalence classes of nominal descriptions based in individuation properties. That is, 'granular aggregates' contains all of the nouns (or noun senses) that encode a construal of entities that includes having perceptible units as well as typically being connected to one another or clumped together.[19] I refer to the scale constructed out of the individuation types, given in 17, as the SCALE OF INDIVIDUATION.[20]

(17) liquids/substances < granular aggregates < collective aggregates < individuals

This view of grammatical countability, where a set of properties bear on lexical classification, is reminiscent of the view of argument realization and thematic roles put forth by Dowty (1991). If, in English, a noun encodes a lexical meaning that rates highly on the various individuation properties, it will be lexicalized as a countable noun, similarly to how if a verbal argument has a large number of PROTO-AGENT entailments and few

[18] The focus on 'perceptible' units here also steers the discussion away from the relation between, for instance, the extension of *water* and individual H_2O molecules, which has sometimes been taken to be a worry in the literature (Quine 1960, Landman 2011).

[19] For a formal semantic analysis of these different individuation types, along with number morphology in Welsh and Dagaare, see Grimm 2012a.

[20] The scale presented in 17 is built from four individuation types, although we will see that the scale of individuation is potentially more fine grained. See Grimm 2012c for discussion of prior formulations of animacy (Lucy 1992) and individuation continua (Gentner & Boroditsky 2001).

PROTO-PATIENT entailments, it will be lexicalized as the verb's subject. While the individuation properties, such as likelihood of contiguous elements, are not entailments active in the on-line calculation of truth conditions, the view on grammatical countability argued for here aligns with Dowty's (1991) PROTO-PROPERTIES to the extent that the proto-properties are taken to be concept-defining features over verbal predicates.

However, individuation properties contrast with the more familiar mereology-based properties, such as divisiveness or cumulativity, which apply conditions at the level of referents. Individuation properties instead delineate information about the conceptualization underpinning a noun, applying at the level of nominal kinds, which then has grammatical effects, such as determining lexicalization patterns. For instance, the knowledge that *rice* refers to a type of entity for which the individual elements typically come in contiguous bunches is an instance of what I term an INDIVIDUATION PROPERTY and is knowledge presupposed about the kind *rice*, but it is not directly entailed or presupposed to hold of the referents designated by the use of *rice* in the way the property of divisiveness typically is—the statement *rice is on the table* may be true even if the elements of rice are scattered and not contiguous. At the same time, I assume that a noun's membership in a particular individuation type determines the sorts of entities that may be denoted by the predicate; namely, if a noun is of an aggregate type, it will be able to denote minimal individual units as well as plural sums or contiguous clusters thereof, while if a noun is of the liquids/substances individuation type, it will not be able to denote minimal units.

I now turn to applying the scale of individuation in 17 to the grammatical number systems discussed so far. I argue that the scale of individuation just developed underlies the structure of grammatical number systems. Laying out the interaction between individuation types and coding types across the different languages of interest here demonstrates that grammatical number systems can be viewed as coherent systems whose morphosyntactic operations respond to the lexical semantic needs of different individuation types.

3.3. RELATING ENTITY TYPES AND THE SCALE OF INDIVIDUATION. I first relate the scale of entity types developed in §2.5, repeated in 18, to the scale of individuation, repeated in 19, clarifying the relation between entity type and individuation type.

(18) liquids < foodstuffs < granular aggregates < vegetation/cereals/fruits ≤ insects < small animals < paired/grouped body parts ≤ middle-sized animals < types of people < individuals

(19) liquids/substances < granular aggregates < collective aggregates < individuals

The scale of entity types, while based on the morphological coding across the different languages examined, fails to be predictive of the structure of grammatical number systems. This was seen for Welsh, where, for instance, not all nouns designating small-sized animals fall under the collective class, but only those whose 'characteristic way of living together [is] in swarms, flocks, herds and shoals' (Stolz 2001:65). In contrast, the individuation properties, such as habitually appearing in groups, cross-cut entity types. Thus, it is the behavior of entities with respect to individuation properties that provides generalizations about the relationship between meaning and grammatical number categories.

While there is no direct relation between the scale of entity types and the scale of individuation, the scale of entity types does contain information about likely members of individuation types in a given language. For instance, if a language treats middle-sized animals as falling into the collective class, then it is likely that it will also treat small animals as falling into the collective class.

From the available comparative data, it appears that languages are comparable in how entity types are classified under individuation types, such as large animals being classified as individuals and cereals being classified as granular aggregates, no doubt due to the causal regularity of the world (see Davidson 1973). Yet there is much room for disagreement among languages and cultures as to whether an entity should be described in one fashion or another. Future typological work may be able to determine whether crosslinguistic variation in grammatical number classification of different entity types—that is, whether the entity types are more or less countable—reliably correlates with differences in construals of the entities as more or less individuated in those language communities.

3.4. RELATING INDIVIDUATION TYPES, GRAMMATICAL CATEGORIES OF NUMBER, AND MORPHOLOGICAL CODING. Having proposed that individuation types are the proper level from which to examine the underpinnings of grammatical number systems, I examine the relation among individuation type, grammatical categories of number, and the manner of morphological coding across the languages under discussion. The view developed here implies a relation among these different elements: a grammatical number category is established based on patterns of morphosyntactic coding, yet the morphosyntactic-coding pattern of a noun, such as how it codes singular or plural, relies on its degree of individuation, that is, its individuation type.

Table 20 compares the systems of Dagaare, Welsh, and English in terms of the scale of individuation, where Welsh represents the tripartite systems discussed in §2. Table 20 displays the individuation types, ordered according to the scale of individuation, along the top. The language is named in the leftmost column. The table displays each of the grammatical categories in the different languages as a shaded region spanning the relevant individuation types. Each of the shaded regions also indicates the grammatical category's manner of coding, for instance for Welsh, (i) no number coding (Ø), (ii) zero-coded nonunit interpretation with a coded unit (*-yn*), or (iii) zero-coded unit reading with a coded plural (*-od*). For Dagaare, the four individuation types are each manifested by a distinct grammatical category of number.

LANGUAGE	LIQUIDS/ SUBSTANCES	GRANULAR AGGREGATES	COLLECTIVE AGGREGATES	INDIVIDUAL ENTITIES
Dagaare	0	0/singulative (*-ruu*)	0/singular (*-ri*)	0/plural (*-ri*)
Welsh	0	0/singulative (*-yn*)		0/plural (*-od*)
English	0		0/plural (*-s*)	

TABLE 20. The scale of individuation: Dagaare, Welsh, and English.

The general picture implied by the table is that morphosyntactic classes are based in the scale of individuation. Each of these systems differs substantially from the others, yet they are still all comparable: they all partition the scale of individuation into a discrete set of grammatical categories. On this view, morphosyntactic categories of number are grammatical phenomena—different languages have established different categories. Yet these categories are based in individuation types, or coherent combinations of individuation types, and therefore are also grounded in the meaning of nouns. Languages may carve up the space defined by the scale in different fashions, but all of these languages are responding to a similar functional need: to provide means of quantifying individuated entities, while not being burdened with count morphology for nouns that are not (sufficiently) individuated.

Recalling that the individuation types are a partitioning of nominal descriptions into their equivalence classes defined by individuation properties, and since morphosyntac-

tic categories of number are also a partitioning of nominal descriptions, one can consider the relation between the two partitions. Although the relation between the cells of the two partitionings could be arbitrary, I argue that morphosyntactic classes adhere to the ordering of the scale of individuation, which I implement by requiring an order-preserving mapping between the cells of two different partitionings.

To put this problem in more precise, and more general, terms, let the set of nominal descriptions in a language be partitioned into a set of individuation types, $\Pi_I = I_1, I_2, \dots I_n$, and also into a set of (language-particular) grammatical number categories, $\Pi_G = G_1, G_2, \dots G_n$. An order-preserving mapping between two arbitrary ordered sets, for example $(P, \leq)$ and $(Q, \leq)$, is a function f such that for all $p_1, p_2 \in P$, if $p_1 \leq p_2$, then $f(p_1) \leq f(p_2)$. In our case, let $Ind = (I, \leq_I)$, the set of individuation types with the ordering given by the scale, and let $Gram = (G, \leq_G)$, the set of grammatical countability classes, also with an ordering on them that will be subject to the order-preserving mapping. Then, any mapping $f : Ind \rightarrow Gram$ is such that if $I_i \leq_I I_j$, then $f(I_i) \leq_G f(I_j)$. In prose, if two individuation types, say granular aggregates and collective aggregates, stand in a less-than-or-equal-to relationship with respect to the scale of individuation, they will map to categories in the language's grammatical-class inventory that preserve the ordering.

To see how this works, consider Welsh. Here f(liquids/substances) = noncountable, f(individual entities) = singular/plural, and f(granular aggregates) = f(collective aggregates) = collective/singulative. The following implications hold: liquids/substances $\leq$ granular aggregates $\Rightarrow$ f(liquids/substances) $\leq f$(granular aggregates) (since noncountable < collective/singulative), while granular aggregates $\leq$ collective aggregates $\Rightarrow$ f(granular aggregates) $\leq f$(collective aggregates) (since collective/singulative = collective/singulative). It can be checked that the other ordering relations are also preserved.

Note that different cells of the Π_I partition could map to a single cell in the Π_G partition, as is the case with the mapping in Welsh from granular aggregates and collective aggregates to collective/singulative; therefore the mapping is not injective. Yet, from the construction of the problem, it follows that it will be surjective—that is, every cell in the Π_G partition will be mapped to from the Π_I partition.

The requirement that grammatical categories of countability respect the ordering of the scale of individuation enforces that grammatical number systems should manifest certain properties, which also has typological implications. First, since grammatical number categories respect the scale of individuation, then, up to equivalence, grammatical number categories will correspond to either individuation types or coherent combinations of individuation types. Second, since the order-preserving map is a function, each cell in Π_I will map to exactly one and only one cell in Π_G. Accordingly, no individuation type will map to two grammatical categories, from which it follows that grammatical number partitioning will be either equivalent to or more coarse grained than the individuation-type partitioning. Finally, since the ordering is preserved, there should not be systems where a category of grammatical number (such as dual or collective or count) spans two disconnected segments of the scale. An example of a system that would violate this condition is shown in Table 21, where individuals and granular aggregates both belong to a singular/plural class, while collective aggregates form a distinct class, which results in the singular/plural class being discontinuous along the scale of individuation.[21]

[21] Here, f(granular aggregates) = f(individual entities) = singular/plural and f(collective aggregates) = collective/singulative. Although granular aggregates < collective aggregates < individual entities, the grammatical categories collective/singulative and singular/plural will not preserve the ordering, since either f(granular aggregates) $\not\leq f$(collective aggregates) or f(collective aggregates) $\not\leq f$(individual entities).

LANGUAGE	LIQUIDS/ SUBSTANCES	GRANULAR AGGREGATES	COLLECTIVE AGGREGATES	INDIVIDUAL ENTITIES
Bad system	0	0/plural	0/singulative	0/plural

TABLE 21. Impossible grammatical number system violating the prediction of the scale of individuation.

The theory here does not provide direct predictions concerning, say, which categories of grammatical number a language will develop, but rather provides broad constraints governing how grammatical number systems behave. These predictions are refined in §4.2 below when the influence of animacy is also considered.

A further prediction concerns the coding of the different number categories, or markedness patterns. There is a choice to be made as to which countability value is the zero-coded, or unmarked, value. For instance, for countable nouns in English the zero-coded value is the singular, while the coded, or marked, value is the plural. In contrast, for the collective class in Turkana, the plural is the zero-coded value, while the singular is the coded or marked value. I suggest that understanding grammatical number categories in light of the scale of individuation provides a basis for the coding choices made in different languages.

Following Jakobson (1971 [1957]) and Horn (1989), I take morphological markedness to be rooted in semantic markedness (see Koontz-Garboden 2007 and Kiparsky & Tonhauser 2012 for further discussion). In particular, I assume that a grammatical number category's coding preferences are linked to the semantic markedness patterns of its associated individuation type(s), and that the default coding value of the grammatical category aligns with the default semantic value of the individuation type. For instance, the individuation type individual entities is composed of things in the world that have a high propensity to appear as free-standing individuals; thus a grammatical class that covers individual entities would likely designate singular entities by default as opposed to multiple entities. That is, the plural would be the marked value, as has been seen in the grammatical number systems so far. Conversely, the individuation type collective aggregates is composed of entities where the default is cooccurrence. For a grammatical class that covers the individuation type collective aggregates, reference to multiple entities, the plural value, would be default, while the singular value would be marked. The limiting case is the portion of the scale for which there is no countability contrast. In general terms, the empirical prediction is that the higher the level of individuation of a grammatical class, the more likely the designation of single entities will be the default, while the lower the level of individuation of a grammatical class, the more likely the designation of multiple entities will be the default.[22] I return to these issues in §4.4.

4. APPLICATIONS OF THE SCALE OF INDIVIDUATION. The scale of individuation as presented accounts for the number systems in §2, which motivated it. This section demonstrates its wide empirical reach, discussing its application to other grammatical number systems, its interaction with animacy, its application beyond natural concrete entities to artifactual and abstract nouns, and its relation to frequency.

4.1. EXTENSION TO OTHER TYPES OF NUMBER SYSTEMS. I now show how the scale of individuation aligns with several other nonbinary grammatical number systems beyond those discussed in §2. While these systems have distinct means of manifesting gram-

[22] This formulation generalizes the notion of LOCAL MARKEDNESS due to Tiersma (1982). See Stebbins 2007 for relevant discussion.

matical number, such as classifiers, the categories of nouns that can be distinguished accord with the scale of individuation.[23]

MIRAÑA. Miraña, a highly endangered Amazonian language spoken in southern Colombia, possesses both nominal classifiers and inflectional number. Seifart (2009) discusses distributional properties that distinguish the different noun types. First, whether a noun combines with inflectional number markers distinguishes countable from noncountable nouns: countable nouns, when designating multiple entities, obligatorily inflect, while noncountable nouns do not, shown in Table 22.

COUNTABILITY CATEGORY	BASE FORM	PLURAL INFLECTED FORM	GLOSS
countable	kaːni	káːni-mɯ	'father'
	ɨːβa	ɨːβa-mɯ	'macaw'
noncountable	kaː	*ka-mɯ	'ants'
	íːnɯ	*íːnɯ-ːnɛ	'earth'

TABLE 22. Countable and noncountable nouns in Miraña.

Noncountable nouns instead combine with class markers to refer to single objects, and these forms then permit inflectional number (as well as cardinal modification), shown in 20–21 (SCM: specific class marker).

(20) a. íːnɯ
earth
'earth'

b. íːnɯ-ba
earth-SCM.3DIM
'piece of earth'

c. íːnɯ-bá-ːkɯ
earth-SCM.3DIM-PL
'pieces of earth'

(21) a. kaː
ant
'ants'

b. ka-ʔba
ant-SCM.3DIM
'an ant'

c. ka-ʔbá-mɯ
ant-SCM.3DIM-PL
'some ants'

There is a distinct class of grammatically noncountable nouns that in their bare form designate a collection. Seifart observes that many nouns, mainly animate but not human, differ from nouns designating inanimate objects or stuff: 'even though the nonunitized forms do not take number inflection, the singular-plural distinction is relevant for them in the sense that the underived form refers to groups of animals and the unitized form to single animals' (2009:39). This can be seen with the word for *ant* in 21, where the bare form refers to groups of ants, and the application of a class marker results in reference to single ants. The plural form, according to Seifart, refers to smaller numbers of entities, glossed by 'some ants'. Thus, this class of nouns in Miraña with default reference to collections aligns with the collective/singulative classes previously explored. Further, the entity types in Miraña's number categories overlap with what would be expected: the aggregate class includes insects along with lower animates, such as rats, whereas the singular/plural class includes humans and big or culturally salient animals, such as the macaw.

In sum, despite the different grammatical means employed, Miraña's classes of noncountable, aggregate, and countable nouns align with the predictions of the scale of individuation, shown in Table 23.

YUDJA. Another Amazonian language, Yudja (Juruna family, Tupi stock, spoken in Brazil), discussed by Lima (2014), serves as a limiting case: the countable/noncount-

[23] See further, for instance, Mathieu 2012 for an application of the scale of individuation to Ojibwe. Other languages for which the predictions of the scale appear to be borne out include Arabic (Acquaviva 2008, Mathieu 2012), Breton (Press 2009), Czech, Hebrew (Doron & Müller 2013), Lingala (Mufwene 1980), Swahili (Contini-Morava 2000), Uncunwee (Jade Comfort, p.c.), and Walman (Matthew Dryer, p.c.).

LANGUAGE	LIQUIDS/ SUBSTANCES	GRANULAR AGGREGATES	COLLECTIVE AGGREGATES	INDIVIDUAL ENTITIES
Miraña	0		0	0/(*-mɯ*)

TABLE 23. The grammatical number categories of Miraña along the scale of individuation.

able distinction is only very weakly present. Nouns in Yudja are allowed to appear bare, unspecified for number or definiteness, as shown in 22, from Lima 2014:33.

(22) ali ba'ï ixu
child paca eat
'The/a/child(ren) eat(s)/ate the/a paca(s).' (lit. 'An undefined number of children eat(s)/ate an undefined number of pacas.')

Lima (2014) reports that the distribution of numeral modification does not distinguish between countable and noncountable nouns: numerals combine freely with nouns designating substances or individuals. Examples in 23 and 24 (Lima's exs. 31c and 28b) are reported as acceptable either in conventionalized contexts, such as containers of sand/blood, or in ad hoc, unconventional contexts, such as clumps of sand or drops of blood. Lima (2014:x) proposes that in Yudja 'all nouns can be used as count nouns'.

(23) Yauda ali eta apapa.
two child sand drop.REDUPL
'Children dropped two (portions of) sand(s) (in different events).'

(24) Txabïu apeta ipide pepepe.
three blood on.the.floor to.drip.REDUPL
'Three (drops of) blood dripped on the floor (in different events).'

The sole discriminating piece of number morphology is an optional plural morpheme *-i* that is restricted to human nouns (Fargetti 2001). Lima further notes that when a human noun has plural reference, use of *-i* is preferred. Clearly, Yudja provides an extreme case where the grammatical traits of a countable/noncountable distinction are hardly present. This system is, however, still consistent with the scale of individuation, as shown in Table 24: only very highly individuated entities (humans) manifest grammatical number, while everything lower on the scale is unspecified.

LANGUAGE	LIQUIDS/ SUBSTANCES	GRANULAR AGGREGATES	COLLECTIVE AGGREGATES	INDIVIDUAL ENTITIES
Yudja		0		0/(*-i*)

TABLE 24. The grammatical number categories of Yudja along the scale of individuation.

KIOWA. The application of the scale of individuation extends to much more complicated systems, such as that of Kiowa (Kiowa-Tanoan, spoken in Oklahoma). Like Dagaare, Kiowa has an inverse number-marking system, but also makes a three-way distinction between singular, dual, and plural number values. The complexity of the system prevents giving a thorough treatment here, but I simply show how the different classes of grammatical number correspond to natural lexical semantic classes in a fashion consistent with the scale of individuation.

Harbour (2008), expanding upon Watkins 1984, establishes nine classes of nouns, including five major ones: (i) animates or inanimates that are capable of motion; (ii) 'plants and plant material, natural and man-made objects and a small number of body parts' (Watkins 1984:85); (iii) vegetation forming natural collections or other nouns disposed toward referring to a cohesive group; (iv) pluralia tantum, abstract nouns, and (for many speakers) granular aggregates; and (v) substances such as milk or honey. There are many interesting nuances, such as the fact that many nouns belong to both

class (ii) and (iii), which then highlight distributive and collective readings, respectively. For the moment, it suffices to note that the major classes align well with what the scale of individuation would predict—ranging from highly individuated entities (class (i)) to vegetation and general objects (class (ii)) to collectives (class (iii)) to pluralia tantum and granular aggregates (class (iv)), and finally to substances (class (v)). To fully connect this system with the scale of individuation, the scale needs to be able to represent both animacy and the dual. I treat animacy in §4.2, but leave the dual for future work since a comprehensive treatment would take us too far afield.

SUMMARY. The different languages examined here exhibit vastly different grammatical number systems in terms of their morphology and degree of complexity. Yet each can be seen as reflecting the organization of the scale of individuation in their own way. It is worth contrasting these results with a view that proposes that countability reduces to a binary ±INDIVIDUAL distinction. Such a view would confront difficulties in the face of Kiowa's wealth of classes, as well as be forced to conclude that what speakers of Yudja consider to be individuals is very restricted compared with what speakers of other languages do. Treating individuation as a scalar phenomenon avoids these difficulties and, instead, speaks to the richness of different number systems.[24] Further, since the notion of individual is a graded concept, it is expected that some languages, like Yudja, might restrict their expression of grammatical number to only the types of individuated entities highest in the scale. I now turn to exploring the relation between the scale of individuation and animacy, followed by its relation with frequency.

4.2. COUNTABILITY AND ANIMACY. I have been arguing that a scalar structure based on individuation underlies countability phenomena; however, other factors, especially animacy, have been shown to influence number marking. While animacy has been reliably tied to the manifestation of plural marking in a large number of languages, the relation between animacy and other types of grammatical number marking, such as the collective/singulative or the dual, has proved to be something of a puzzle. I show that viewing grammatical number as grounded in individuation leads to understanding how animacy influences the inventory of nouns that fall into the collective/singulative category in a given language.

The pioneering typological study by Smith-Stark (1974) demonstrates that the degree of animacy of a noun's referent is correlated to the likelihood of that noun being able to express plural marking. Smith-Stark (1974) develops the scale in 25, which he considers a scale of the likelihood of participation in the speech event. The associated claim is that if a language expresses plurality for noun types in a given position on the hierarchy, then the language will also express plurality for any noun type higher on the hierarchy.

[24] A referee remarks that distinguishing the singular/plural contrast from the mass/count contrast is sufficient to explain phenomena like plural mass nouns (*clothes*, *dregs*) and collective nouns, citing Gillon 1992 as an example for which a feature system is constructed from two equipollent syntactic features, ±Count and ±PL, resulting in three categories: count nouns (+Count) (which may be singular or plural), singular mass nouns (−Count, −PL), and plural mass nouns that are −Count but lexically specified as +PL. Despite the fact that such systems are richer than a strictly binary ±Count, they do not provide a rich enough framework to treat the range of phenomena discussed here. Even the variation in collective/singulative systems discussed in §2 poses difficulties. Collectives would be relegated to the third category, −Count, +PL, which could, at a first pass, function for Welsh and Turkana, whose collective forms impose plural agreement, but not for Maltese, where the collective agrees in the singular. Such a treatment would further remain silent on a range of issues, for instance on interpretational differences between the collective classes discussed here and plural mass terms (*brains*) or on grammatical number systems whose complexity exceeds three categories, as for Dagaare or Kiowa.

(25) The Smith-Stark hierarchy:
inanimate < animate < human < rational < kin < addressee < speaker

An instantiation of this prediction is provided by Tamil, as shown in 26. The shaded area of the animacy hierarchy in 26 indicates the portion that manifests a singular/plural contrast. According to the discussion in Smith-Stark 1974:662, nouns in Tamil designating rational beings typically manifest plurality, while those designating living beings not considered rational, including *child* and *infant*, are rarely pluralized. Smith-Stark's claim is that any noun type higher on the hierarchy would also have obligatory number marking, which is true—for example, all pronouns manifest a singular/plural distinction.

(26) Tamil's grammatical number system along the Smith-Stark hierarchy:
inanimate < animate < human < rational < kin < addressee < speaker

While the hierarchy in Smith-Stark 1974 has proven remarkably robust for predicting the occurrence of plural marking, its relation to other grammatical number categories has remained opaque. Corbett (1996) discusses data pertaining to what he terms MINOR NUMBER categories—essentially, any number category that is not the singular/plural category, such as dual, paucal, collective, or mass. Corbett (1996) provides data from eight different languages that prove problematic for the Smith-Stark hierarchy. In each case, the number categories, such as the collective, are apparently not sensitive to the animacy level of the noun.

For instance, Corbett points out that the collective in Maltese contains a mixture of animate ('flies') and inanimate ('corn', 'shoes') nouns, but cannot be expressed on, for instance, first- and second-person pronouns, violating the predictions of the Smith-Stark hierarchy. This is shown in 27, where the shaded region indicates for which part of the hierarchy the collective/singulative is expressed.

(27) Smith-Stark hierarchy violated by Maltese collective:
inanimate < animate < human < rational < kin < addressee < speaker

The question arises, then, of what precisely the relation is between grammatical number categories such as the collective and the animacy hierarchy. From the data considered so far it would seem that the two are independent—the use and function of a grammatical number category such as the collective is picking up on attributes of entities that are not directly related to animacy. This is the conclusion that Corbett comes to: while the Smith-Stark hierarchy governs the portion of the lexicon to which grammatical number generally applies, there may be 'patches' of the grammar where minor number applies. In these cases, a very particular semantics is at play. Indeed, for some of the examples Corbett considers, such as the collective plural in Budugh (Lezgian, spoken in northeastern Azerbaijan), which has a very limited distribution, being restricted to just five nouns referring to paired body parts, much more does not need to be said. Yet, for the systems considered in §2, which have smaller categories of number that are nonetheless productive, one would hope something more general could be said about the relation between number categories, individuation types, and the animacy hierarchy.

Viewing grammatical number as related to individuation provides the key to understanding the behavior of the collective/singulative class in relation to the animacy hierarchy. In order for nouns to be members of the collective/singulative class, the tendency to view the entity as habitually occurring in a group must be greater than the tendency to view the entity as habitually occurring singly. Yet the higher one ascends on the animacy scale, the more individuated the entities are, and the greater the tendency to view them as occurring singly as opposed to coming in groups. Thus, from the viewpoint of individuation, the occurrence of the collective/singulative category should be IN-

VERSELY related to the animacy hierarchy. In other words, if a language possesses a collective/singulative class, the higher a noun rates on the animacy hierarchy, the less likely it will fall into the collective/singulative class. This is the inverse of the claim Smith-Stark (1974) makes for plural marking.

This discussion can be put in a more precise form by representing the different possibilities that arise when the scale of individuation interacts with the animacy hierarchy. A simple technique for multiplying different linear scales was demonstrated by Aissen (2003), and I refer to it in what follows as an AISSEN LATTICE. Here it is necessary to take the product of the countability scale and animacy hierarchy. I modify the animacy hierarchy from the form given by Smith-Stark (1974). First, I do not represent the categories above HUMAN, as they are not relevant for the distribution of the collective. Second, it has been noted that many languages do not treat all members of the category ANIMATE equally in terms of number marking. Haspelmath (2005) notes that the category of animates is often subcategorized into 'higher' and 'lower' animates, which I adopt here. This distinction was already seen in Miraña, which categorizes salient animates, such as those of high cultural value ('macaw'), with humans in the singular/plural class, as opposed to other animates whose nouns in their base form had collective reference.

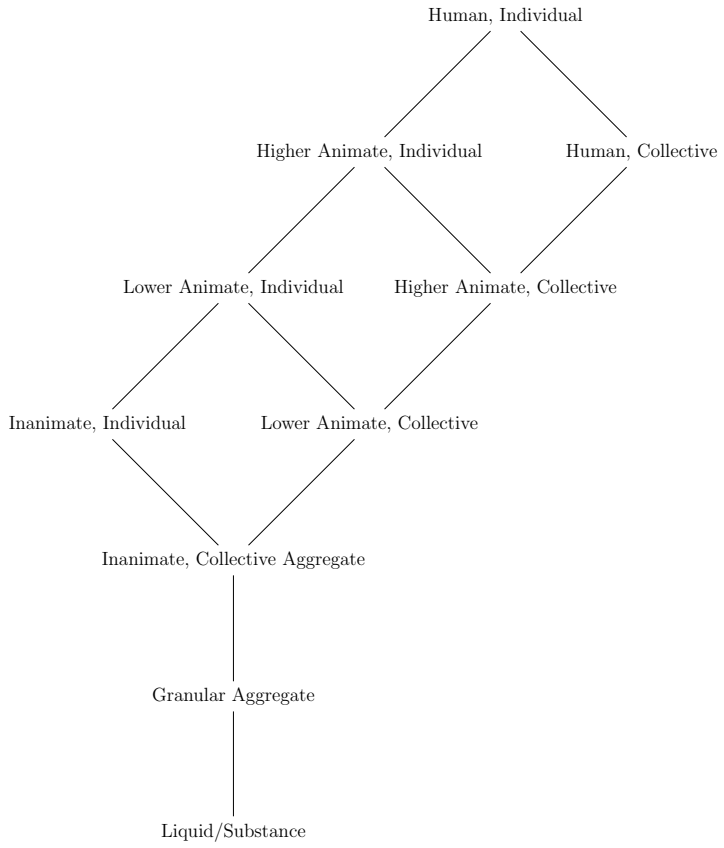

FIGURE 3. The lattice of animacy and individuation.

The Aissen lattice in Figure 3 exhausts the combinatoric space of the individuation scale and the animacy hierarchy. For the segment of the individuation scale from collective aggregate and higher, the entirety of the (simplified) animacy hierarchy is attached

to each node. I consider granular aggregate and liquid/substance to be necessarily inanimate, and therefore incompatible with the rest of the animacy hierarchy, so they are just represented by single nodes.

In Figure 4, the systems of Welsh, Turkana, and Maltese are represented on the combined animacy-individuation lattice. The extent of the collective/singulative class in each language is indicated by shading, which covers the relevant nodes of the lattice.[25] While these three languages are highly similar at the abstract level of the individuation hierarchy, through this mapping we can see differences across the different systems.

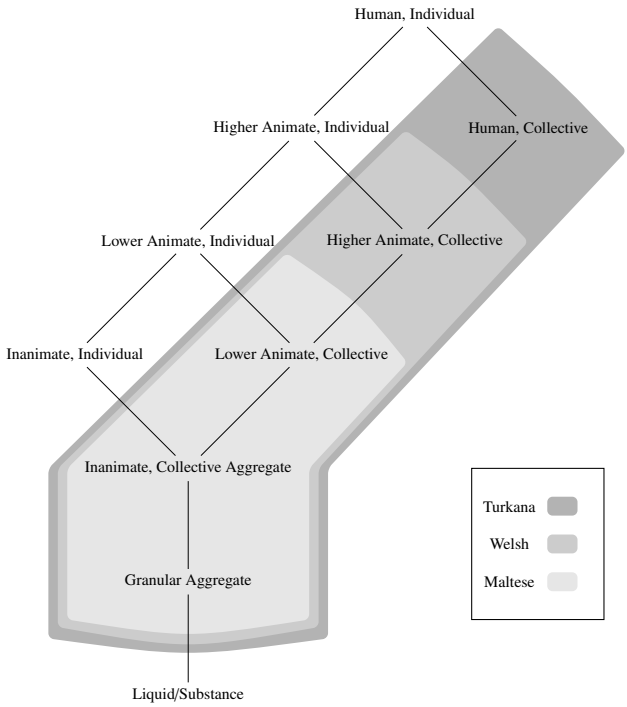

FIGURE 4. The collective/singulative classes in Turkana, Welsh, and Maltese on the lattice of animacy and individuation.

Returning to the relation between the animacy hierarchy and minor number categories, it appears that there is an interaction between the distribution of the collective categories in these languages and the animacy hierarchy. Yet, unlike plural-marking across languages, which differs in how far it descends the animacy hierarchy, the collective/singulative classes in these languages differ in how far they ASCEND the animacy hierarchy. For instance, the Welsh collective/singulative class contains inanimate nouns, insects, and small and middle-sized animals. Turkana's contains those and in addition human nouns; thus, it includes a larger but still connected section of the hierarchy.

The collective/singulative category in Maltese, however, spans a smaller segment of the animacy hierarchy. Maltese has a restricted number of animate members in its collective/unit category, essentially limited to insects. This restriction is somewhat surprising given that this category historically developed from the Arabic collective, which, like Turkana's, contained nouns designating larger animals, such as cows, and collec-

[25] I simplify by leaving aside the packaging uses of Maltese, as this is not relevant for the moment.

tions of humans, such as names of nationalities or professions. Mifsud (1996) notes that there has been a historical shift whereby in Maltese these nouns were integrated into the singular/plural category. The important point for the moment is that this decrease in membership seems again to proceed along the lines of the animacy hierarchy. What does not appear to be attested is a collective/singulative class that includes many nouns referring to humans and many referring to inanimates, but none to other animates—that is, one that would be discontinuous.

Given this evidence, the conjecture about the structure of grammatical number categories can be refined with respect to animacy. This principle, the CONNECTED REGION CONSTRAINT ON GRAMMATICAL NUMBER CATEGORIES, is stated in 28. That is, a grammatical number category should not be discontinuous across either segments of the individuation scale or segments of the animacy hierarchy.

(28) CONNECTED REGION CONSTRAINT ON GRAMMATICAL NUMBER CATEGORIES: Grammatical number categories occupy connected regions of the animacy-individuation lattice.

From this perspective, the minor number categories do actually respect the animacy hierarchy, but simply in a different fashion than plural marking does. The singular/plural distinction occupies the upper regions of the animacy hierarchy, and any extension into the lower categories is predicted to occur in a continuous fashion. The collective/unit distinction occupies the middle region of the animacy hierarchy, and any extension into the upper/lower categories is predicted to occur in a continuous fashion. The functional grounding for this division of labor is evident: higher animate entities, and certainly speech-act participants, are clearly individuated. Thus, if anything will be high on the individuation hierarchy, it will be such entities.

4.3. BEYOND NATURAL CONCRETE ENTITIES. The proposed view extends to other types of nouns, such as artifactual and abstract nouns. On the view here, such nouns are also categorized into grammatical countability classes based on individuation properties, yet the nature and kind of individuation properties are different. While perceptual properties (shape, spatial proximity) are relevant for individuating natural concrete entities, they will not be relevant for individuating, say, eventive nouns, where properties such as temporal discreteness or continuity of the event described are likely to be relevant (Mourelatos 1978, Bloom 1990). In this section, I sketch how artifactual nouns, such as *chair* or *furniture*, can be treated in the view developed here.[26] *Furniture* has long been regarded as paradoxical: it is noncountable, yet unlike noncountable substance nouns, *furniture* designates individual objects, such as chairs or tables, which are otherwise countable. Certain linguistic contexts demonstrate that the individual objects are grammatically accessible, as with comparatives (Barner & Snedeker 2005) or adjectives like *big*, as in *big furniture* (Schwarzschild 2011). If countability only relates to a binary ±INDIVIDUAL distinction, then the behavior of *furniture* is unexpected as it is grammatically noncountable despite being capable of referring to individuals.

The behavior of *furniture* would be similarly unexpected if individuation could only be achieved through perceptual properties. Yet artifactual nouns differ from natural concrete nouns in that they do not describe merely objects but objects with respect to a function.[27] I argue that individuation for artifactual nouns is bound up with the function described by the noun.

[26] See Grimm & Levin 2017 for a lexical and formal semantic analysis of artifactual nouns.

[27] This has been widely noted since Aristotle. See Brown 1999 and Nichols 2008, among others, for linguistic effects of this distinction.

The functions associated with artifactual nouns further provide a contrast between individuals and aggregates of individuals, from which an explanation of *furniture*'s countability classification follows. The functions associated with *chair* or *hammer*, namely to sit on or to hammer with, are canonically achieved through using just one entity. In contrast, the function associated with *furniture* is to furnish a space, as indicated by its derivational source, *to furnish* (> *fournir* Fr.). Furnishing a space is canonically achieved through using multiple entities. Thus, an individual/aggregate distinction emerges for artifactual nouns as well: nouns such as *chair* describe entities that individually are capable of performing the associated function, while nouns such as *furniture* describe entities that typically perform the associated function as an aggregate. Accordingly, *furniture* is classified in English in the same individuation type as other aggregates where access to individuals is not relevant to their description, such as *foliage* or other granular aggregates. Since aggregate nouns such as *foliage* are treated as noncountable in English, *furniture* is also classified as noncountable. An immediate prediction is that cognates of *furniture* may appear in a collective/singulative class should the language have one. This prediction is borne out by Welsh, where the cognate for *furniture* appears in the collective/unit class, *dodrefn* 'furniture'/*dodrefn-yn* 'a piece of furniture', as do other functional aggregates such as *offer* 'tools'/*offer-yn* 'a tool' or *dillad* 'clothes'/*dilled-yn* 'a piece of clothing'. Thus, once distinct principles of individuation are recognized for natural concrete and artifactual nouns, then nouns such as *furniture* need not be seen as an exception to the principles of countability classification, but as aligned with them.

Placing the explanatory burden on domain-specific modes of individuation also yields a more nuanced understanding of countability preferences: although *oil* and *furniture* may both be noncountable in English, they need not be so for the same reason. Rather, their respective grammatical classifications as noncountable nouns derive from distinct modes of individuation, related to their differing lexical semantics. This differs from accounts that give a uniform explanation for the noncountable status of *oil* and *furniture* in English (Chierchia 1998), or where *furniture* is treated as a lexically specified exception (Bale & Barner 2009). The countability behavior of artifactual nouns, such as those lexicalizing a description of furniture, will systematically differ across languages according to which type of function is described by the noun, its related individuation properties, and the grammatical system of the language at hand. The resultant countability status may coincide with that of substance nouns, for example, *oil*, but often will be distinct. This view anticipates crosslinguistic differences exactly when the nominal description differs. Cognates of *furniture*, such as *meuble* in French, are countable, which has been taken as problematic. Yet this is consistent with the approach advocated here since *furniture* and *meuble* differ at the level of entity description: *meuble* is derived from a root meaning 'to move' and thus designates 'movable entity', which in turn evokes different individuation properties from entities related to the root *to furnish*.

Applying this discussion to the scale of individuation indicates that ultimately the units of the scale must be generalized, composite types. Clearly, there will be multiple, domain-dependent modes of individuating entities, including spatial separation, temporal separation, or individuation via function. The 'individuals' or 'collective aggregates' individuation types will contain entities whose individuation properties differ in their domain-specific characteristics, such as being spatially separated versus temporally separated, but that share in abstract individuation properties, namely being separated. It is incumbent on this view to understand the principles of individuation that underlie nominal categorization for each type of entity and to articulate in detail the relation be-

tween the lexical semantics of a noun and the individuation properties relevant to the noun type. The expectation then is that, modulo the specific facets of the domain-dependent individuation properties, the categorization of natural concrete, artifactual, eventive, and abstract nouns will maintain the higher-level individuation contrasts that ground grammatical classification of countability.

4.4. INDIVIDUATION, MORPHOLOGICAL CODING, AND FREQUENCY. I now consider the relationship between the proposed individuation account and frequency. The singular/plural and collective/singulative classes display coding asymmetries: one form, say the singular, is zero-coded (or unmarked), while the other form, the plural, is overtly coded. Typological studies have sometimes attributed such coding asymmetries to frequency asymmetries (Haspelmath 2008). Two notions of frequency are relevant here. First I discuss nonlinguistic frequency, namely frequency in relation to elements in the external world, and then linguistic, or textual, frequency, that is, frequency in relation to occurrences of words or morphemes. Finally, I discuss the view wherein economy (using less coding for more frequent forms) and learnability are two contrasting forces that shape language systems and how that view aligns with the patterns found in the different grammatical number systems reviewed here.

REFERENCE AND FREQUENCY. It is tempting to think that nominal countability categorization could be explained purely in terms of how frequently entities cooccur in the world, as suggested by Tiersma (1982) or Haspelmath (2008). For instance, it is reasonable to suppose that since ants typically appear in great numbers, it is more economical to use a word form for which the plural interpretation is the default. Yet, from the sources consulted for the various languages examined here, frequency of cooccurrence appears to be, at best, a necessary but not sufficient condition for a noun to fall into a collective/singulative class.

Stolz (2001:65) argues explicitly that for Welsh whether a noun is assigned to the collective/singulative class is not 'a matter of quantity'. In part, it depends on the behavior of the entities; for instance, for animate entities it is 'the characteristic way of living together in swarms, flocks, herds and shoals' (ibid.). Stolz further argues that gregariousness of entities is also not in itself sufficient, providing several counterexamples such as the Welsh term for 'cow', an animal that does come in herds but is a member of the singular/plural class. Rather, it is whether the entities are PERCEIVED as habitually coming together that is determinate, according to Stolz.[28] Similarly, for Turkana, Dimmendaal observes that 'living in herds or groups' (2000:229) is the relevant characteristic for nouns designating animate entities that are zero-coded in the plural. Further, as observed in the discussion of Welsh and Turkana, there is a meaning contrast between collective and plural values, and therefore it seems unlikely that nouns are categorized in the collective/singulative class simply by virtue of the noun typically referring to multiple entities, namely a plural value. Rather, the noun must refer to entities habitually appearing in some number, which additionally must be viewed as linked together, for instance, through collective living, or related to a common source, such as multiple berries related by growing on the same branch of a bush. In sum, if an entity appears in

[28] Stolz (2001) is similarly critical of the factor of size, noting that while there is a high proportion of nouns that designate small entities or animals, this does not cover all of the cases and appears to be a secondary consideration compared to swarming or herding behavior. This aligns with the results of the first experiment of Middleton and colleagues (2004), presented in §3.1, where size was not a significant factor.

multiples, that is a precondition, rather than the ultimate explanation of why a noun may be classified in the collective/singulative class.

TEXTUAL FREQUENCY. Coding asymmetries, and the notion of 'markedness', have been related to textual frequency, as in Greenberg 1966, where zero-coded forms are shown to have greater textual frequency than overtly coded ones. The frequencies relevant here are the frequency with which an entity is spoken of as referring to multiple referents as compared to the frequency with which it is spoken of as referring to a singular referent. Although textual frequency has been causally related to the structuring of morphological coding (Greenberg 1966), the different strains of evidence brought together here, such as the psycholinguistic evidence in §3.1, indicate that countability classification cannot be reduced to textual frequency.

However, I demonstrate that there is a correlation between morphological-coding asymmetries and textual-frequency asymmetries: nouns that are typical members of a morphological-coded collective/singulative class correspond, for languages that manifest only a singular/plural contrast, to nouns that have greater textual frequency in the plural than in the singular.

If the semantic domains relevant to 'unmarked plurals' in languages like Welsh or Dagaare are such that speakers more frequently refer to multiple referents than singular referents, then languages that do not possess morphologically unmarked plurals should still display asymmetries in the same semantic domains in terms of textual frequency. In order to evaluate this prediction, I examined frequencies for English nouns in the semantic domains of animals and insects from the COBUILD corpus (18 million words) provided by CELEX (Baayen et al. 1996). Using basic terms and terms that had correspondents in the vocabulary of Dagaare, I calculated the plural-to-singular ratio for these two domains, shown in Figure 5, where the x-axis represents the ratio of the token frequency of plurals to the token frequency of singulars and the y-axis represents the number of lexical items. The graph indicates that there is a clear trend for insect terms to have a plural/singular ratio greater than 1, indicating that insect terms occur more frequently in the plural, while animal terms tend to have a plural/singular ratio less than 1, indicating that animal terms occur more frequently in the singular.[29]

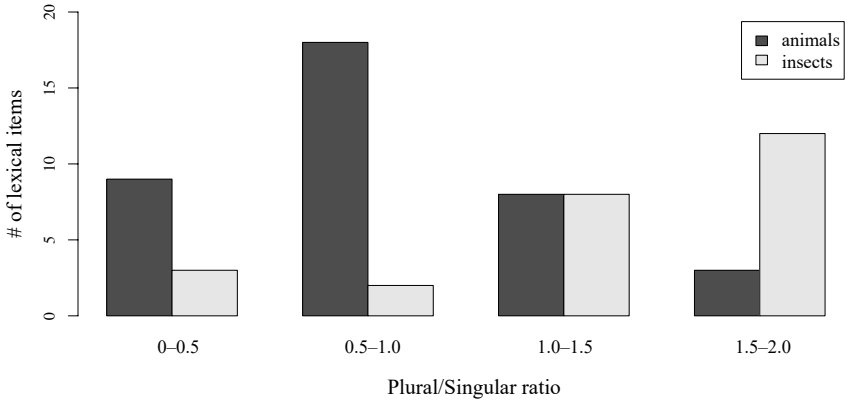

FIGURE 5. Number-marking frequency patterns in English for nouns in the semantic domains of animals and insects.

[29] Similar evidence is provided by Baayen and colleagues (1997), who investigate grammatical number morphology and the phenomenon of local markedness (Tiersma 1982) in Italian from a psycholinguistic perspective.

ECONOMY AND LEARNABILITY. Given the correlation between the entity types that fall into collective/singulative classes and textual frequency, it stands to reason that zero-coded plurals or collective countability classes are beneficial from the point of view of economy of expression.[30] In English, for many words such as the insects in Fig. 5, one must more often than not pronounce an *-s*, an option less economical compared to treating multiple reference as the default. In languages such as Dagaare or Welsh, the default form simply corresponds to the more frequently used form.

The trade-off for increased economy gained through default forms that refer to multiple referents is increased complexity of what must be learned. While in English children must only learn two categories of nouns with respect to number—noncountable nouns and those with a singular/plural contrast—in Welsh or other languages with a collective/singular class, there is an extra category that must be learned. In fact, this appears to require some effort on the part of children acquiring such systems. Although research on the acquisition of such grammatical number systems is still largely unexplored, some results indicate that collective categories are more complex to master. Ravid and Hayek (2003) have investigated the acquisition of collective forms in Palestinian Arabic in children from ages four to eight, where their results indicate that collective nouns in Arabic are quite difficult to acquire. While the oldest children performed with around 85% accuracy for duals and plurals, they only had, at best, 50% accuracy for collective forms.

In sum, while frequency alone is not sufficient to explain the coding patterns of the different grammatical systems discussed here, there is an intimate link between which word forms correspond to the default situation in the world and textual frequency. In turn, this supports a view of languages such as Dagaare as providing an economical grammatical number system. This concludes my discussion of the scale of individuation. I now turn to examining its implications for the theoretical understanding of countability.

5. A MEANING-BASED ACCOUNT OF COUNTABILITY. The previous section has made the case that countability should be understood in light of several interrelated elements: grammatical number categories, individuation properties, and entities in the world. The proposal is represented in Figure 6, which unpacks the four different levels: things in the world, lexical nouns and their properties (that is, entity descriptions as discussed in §3.2), individuation types, and grammatical classes.[31] The figure displays the full generality of arbitrarily many individuation types and grammatical classes, and, to exemplify, mappings are given of two entities that belong to the highest and lowest individuation classes.

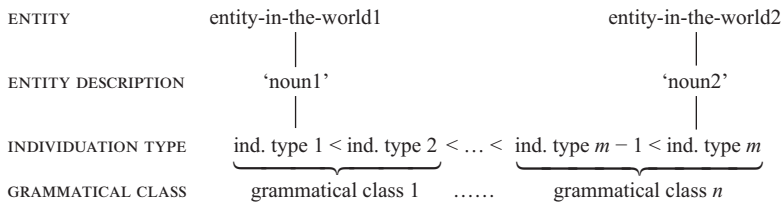

FIGURE 6. Mapping between entities in the world, lexical nouns, individuation types, and grammatical classes.

[30] I would like to thank Paul Kiparsky for leading me to the connections in this section.

[31] This separation of different levels is implicit in much work that emphasizes individuation and is also explicitly laid out in a similar fashion in Bloom 1990:107, although only a binary countability distinction is put forth there.

The level ENTITY—the thing in the world picked out by the noun—contrasts with ENTITY DESCRIPTION, the noun itself as a DESCRIPTION of a type of entity. The entity description encodes a particular construal of the entity, not just the extensional content. Importantly, the entity description may contain entailments specifying which sorts of referents a noun designates. For instance, the entity description in Welsh *cacwn* 'hornet' encodes that these entities may appear as spatially proximate clusters of individual entities or simply as individual entities. Welsh encodes spatially proximate clusters as the default (collective) form, and *cacwn* can be analyzed as entailing that its denotation is restricted to spatially proximate clusters.[32] In contrast, although English speakers have knowledge that hornets often appear in spatially proximate clusters, this is not encoded as a condition on the referents designated by the English noun *hornet*; rather, the English term (in default singular form) entails that its denotation is restricted to individual entities.

Although a distinction between entities and entity descriptions is less frequently made in the countability literature, a distinction between EVENTS and EVENT DESCRIPTIONS has long been recognized as crucial in the aspect/telicity literature (Krifka 1992). The same event can be described by two different expressions that differ in telicity, for instance, *running* (atelic) and *running a mile* (telic). If one assumes that these expressions refer directly to events, then this situation is paradoxical, for the same event is at once atelic and telic. Recognizing that these expressions are simply two different descriptions of the same event immediately resolves this conflict. As I discuss in §5.2 below, recognizing the distinction between entities and entity descriptions similarly resolves paradoxical examples such as *leaves* and *foliage*—both words may refer to the same stuff in the world on a given speech occasion, but they are different descriptions of that stuff.

The proposal in Fig. 6 is a shift in perspective from previous theories of countability. First, there is a higher degree of complexity both in the relation between the things in the world and the grammatical class, represented by the vertical dimension, and also among the individuation classes and the grammatical classes, represented by the horizontal dimension. Second, on this view countability involves a PRINCIPLED RELATION between things in the external world and grammatical classification that is mediated by individuation.

Both directions of this relation are relevant for different facets of countability. One direction represents the world-to-word fit, guiding lexicalization patterns. Things in the external world and how we interact with them set the possible ways the entity may be construed, or individuated. A noun, or in more complicated cases one sense of a noun, lexicalizes one of these construals. The entity description encoded by the noun falls under an individuation type, which determines the noun's countability status with respect to the grammatical number system of the language. The other direction represents the word-to-world fit, guiding how interpretation is achieved. There is a relation between grammatical classification and types of entities in the world and, accordingly, interpreting a noun in tandem with a grammatical classification that specifies particular individuation properties will aid in picking out the intended referent in the world.

Countability does not arise from any single level, but is a product of the interrelation among these different elements. The additional complexity invoked here has a pay-off. Other theories of countability are primarily motivated either by the regularity of the grammatical classification of things in the world into countability classes or by the variability of this classification, or yet again, by the variability possible due to different

[32] See Grimm 2012a for a formal implementation.

contexts. Accordingly, while these theories have the advantage of simplicity, they explain one facet of countability classification at the expense of the others. The complex of relations in Fig. 6 addresses these different aspects of countability simultaneously. I now relate a range of other theories in reference to the different levels given in Fig. 6.

The ontological and grammatical views, introduced in §1, address only part of the picture given in Fig. 6. The ontological view on countability places the explanatory burden of countability on the properties of entities in the external world, the entity level, and presents countability as a function from entities to grammatical classification. Accordingly, for the ontological view, countability involves only two of the levels in Fig. 6 and one direction, as it is a function from the level of entity to grammatical class. The grammatical view, by contrast, places the explanatory burden on the relation between nouns and grammatical class, an ultimately arbitrary relation independent of 'how things are inherently structured'. Accordingly, the grammatical view also involves a relation between only two of the levels in Fig. 6, namely entity description and grammatical class.

While the ontological and grammatical views speak, respectively, to the regularity and variability observed in countability classification, these simpler models of countability run aground on the complexity of the data reviewed here. The exclusion of the level either of entity or of entity description renders many of the phenomena reported mysterious. For instance, meaning shifts, as exemplified by the word *chad* discussed in §3.1, do not accord with either view. For the ontological view, the inherent properties of the entities falling under the noun *chad* have not changed, so no change in countability class should arise. For the grammatical view, the relation to the external world is irrelevant, so no change in countability class should arise either.

A very different set of theories take countability to be a function of (morpho)syntax. Borer (2005), building on Sharvy 1978, argues that nominal roots, the level of entity description in Fig. 6, designate only 'stuff' and contain no information related to countability—only through morphosyntactic structure, such as plural marking or classifiers, can a countable interpretation result. This is the mirror image of the ontological view. There, in order to determine if a noun is countable, one looks to the world, which then determines the grammatical class. Here, in order to determine what a noun refers to, one must look at the grammatical class information, which determines whether the noun refers to discrete entities or stuff. Thus, these structure-based theories provide a radically different view on the locus of countability: it arises solely, or primarily, depending on the variant of the theory, via grammatical structure, and little or no information about countability is contained at the lexical level, that is, at the level of entity description. This is one step further than the grammatical view: now countability is not just independent of 'how things are inherently structured', but truly indifferent to it. Since these theories are motivated by contextual variability of countability, that *chad* has both a countable and noncountable variant poses no difficulty. Yet these theories make strong predictions about the pervasiveness of contextual variability, which I return to in §5.1.

In the remainder of this section, I address how the theory proposed here differs in its predictions from the grammatical, ontological, and syntactic views on countability. I first discuss the challenges from the syntactic approach, since it poses the most dramatic challenge by denying lexical categorization altogether. I discuss its empirical shortcomings and then how contextual variability works within the account here. After arguing that lexical categorization cannot be dispensed with, I turn to examining the different predictions that these accounts make for regularity and variability in countability classification both within and across languages. I demonstrate that lexical and

crosslinguistic variation is not incompatible with a meaning-based theory of countability, but is actually a prediction of this more articulated account.

5.1. RECONCILING LEXICAL MEANING AND CONTEXTUAL VARIABILITY. The model in Fig. 6 speaks to the regularity of the relation between nominal reference and countability classification: a noun encodes an entity description, a particular construal of the entity, which then uniquely determines its countability classification. Yet this would seem to be obviously contradicted by the many nouns that occur in more than one countability category. One set of cases, termed 'flexible nouns' by Barner and Snedeker (2005), concern nouns such as *string* that are regularly used with both countable and noncountable syntax (*three strings*/*a lot of string*). The other set of cases include very many nouns that regularly fall into one countability category but can nevertheless be used in a different countability category when the context is appropriate, as in the well-known cases of grinding, packaging, and the like. I consider these two types of cases in turn under the heading of CONTEXTUAL VARIABILITY.

The existence of contextual variability has been taken by some researchers to vitiate, either partially or totally, the possibility that the lexical meaning of a noun incorporates a predictable relation to a countability categorization. Considering flexible nouns, Barner and Snedeker (2005) assert that there is no lexical specification; rather, for those nouns, countability arises only through the syntactic environment. That many nouns, given an appropriate context, can have both countable and noncountable uses has led to even stronger views, such as that put forth by Borer (2005), where the lexical level plays NO role, and all information about countability is contributed by the syntax. Although providing a highly economical theory, this strong view has suffered a number of empirical setbacks. Bale and Barner (2009) point out that the English noun *furniture* and other 'object-mass' nouns such as *mail* are counterexamples: they must have access to individuals in their denotations despite strongly resisting pluralization or other count morphology. Thus, there must be lexical classification at minimum for such nouns. Other researchers point out that in several languages bare nouns designate individuals rather than 'stuff', despite lacking the syntactic structure that, by hypothesis, would be required, as in Dëne Sųłıné (Wilhelm 2006) and Karitania (Doron & Müller 2013). Dagaare, as has been discussed, is also such a language—the bare use of Dagaare's term for 'dog' is acceptable only as referring to individual dogs, not to dog meat or other interpretations that may be found with noncountable terms in English. In sum, the strong view that there is no lexical classification and that nouns can be interpreted as individuals only through the contribution of syntax appears untenable in light of recent empirical work.

One could consider a mixed view, as in Bale & Barner 2009, where a portion of the lexicon, in particular their object mass nouns (*furniture*), is specified to have only individuals in the denotation, but the rest of the lexicon is underdetermined and differentiated only by syntactic context. This view clearly speaks to the observed VARIABILITY of nominal reference for this larger portion of the lexicon, for there are no constraints put on it. Yet there is a high price attached to this analysis since it does not make predictions for REGULARITY of nominal reference, and so does not connect to the typological generalizations.

It turns out that both sorts of contextual variability are straightforward to accommodate in the model advocated here. Flexible nouns, such as *string* or *chicken*, are simply polysemous between two different entity descriptions. Each of these entity descriptions follows the predictions of the mapping given in Fig. 6 when applied to English: the sense of *chicken* that describes living animals falls into the grammatical class of count-

able nouns, while the sense that describes meat falls into the grammatical class of noncountable nouns. From this perspective, flexible nouns provide little of interest for theories of countability aside from possessing a type of polysemy that straddles two countability classes, while other types of nominal polysemy, for instance between the institution and building senses for *bank*, do not. Some cases that have been treated as grinding (*chicken*) or packaging (*a beer*) fall within this explanation. Thus, these nouns present no particular obstacle to the view that nouns lexicalize countability information—one must only grant that the phenomenon of polysemy exists, which I take to be uncontroversial. This would have an impact on our theories of countability only if flexibility was a general property of nouns—a consideration to which I now turn.

The claim that it is possible to use nouns that to all appearances are countable nouns in noncountable contexts and vice versa has a long history, going as far back as Gleason 1965: 'Every noun, given the right context, can occur in either type of usage, count or mass' (pp. 136–37; see also Pelletier 1975, Allan 1980, Bunt 1985, Pelletier & Schubert 2004, Borer 2005, among others). Should this statement be empirically correct, then a clear solution is to assimilate all nouns to the status of flexible nouns: all nouns will permit some meaning consistent with a noncountable use and some meaning consistent with a countable use. A lexical entry for nouns is then at once impoverished in terms of possessing any sort of countability information, but extremely well-off in terms of access to meanings that are compatible with different countability contexts. A risk for this position, of course, is that it generalizes to the worst case.

I argue that one can accommodate contextual variability while still maintaining a contentful lexical semantics of nouns, namely through pragmatic reasoning underpinned by the model in Fig. 6. Take the utterance 'There is armadillo all over the road'. By hypothesis, there is no single lexical item that corresponds to the meaning of splattered armadillo, and accordingly there is a lexical gap. Using a bare form of *armadillo* serves to fill the lexical gap in this context. Even though neither a speaker nor a hearer has access to a lexicalized bare form of *armadillo* that indicates 'armadillo stuff', the use of *armadillo* given the context is enough for the hearer to work out the intended meaning, of which the speaker is aware, based on analogy with bare nouns that do designate stuff and in the standard way pragmatic models have led us to expect.[33] I now elaborate predictions that this pragmatic-based model makes but that a theory for which all or most nouns are flexible nouns does not.

First, the lexicalized meaning of a noun provides the sine qua non use of the noun, such as *hat* as a clothing item worn on the head, and extralexical uses arise through pragmatic reasoning, such as *hat* as referring to shredded material coming from a hat in sentences like *There is a hat all over the floor*. Accordingly, the extralexical uses of a given noun should be pragmatically marked. Indeed, in contrast to bona fide flexible nouns where both senses occur frequently and unremarkably in countable and noncountable contexts, many of the grinding examples in the literature have a novelty and/or a jocular nature, as noted by Bloom (1990), and may also involve reimaginations of reality, such as the early example from Gleason (1965), given in 29.

(29) Mother termite complains about her son Johnny: 'Johnny is very choosy about his food. He will eat book, but he won't touch shelf.'

These qualities are hallmarks of nonce uses, as opposed to standardized uses, of a word form. Such nonce uses are well known from the literature on language change: novel uses of a word form often serve an immediate communicative need for the speaker,

[33] See also discussion in Falkum 2010 concerning mass-count polysemy and pragmatic reasoning.

which may include 'economizing, filling in a conceptual/lexical gap, or creating a stylistic effect' and which furthermore 'must be understood from its context' (Brinton & Traugott 2005:45). The vast majority of examples of grinding fit this description.

This explanation of contextual variability of countability connects to other widely discussed uses of nouns that are extralexical and whose interpretation is guided by contextually based pragmatic reasoning, such as in *The ham sandwich wants his check* (Nunberg 1979). For these cases as well, there is a stylistic effect achieved (see Nunberg 2004 on 'noteworthiness' as a condition on predicate transfers). It is telling that for these cases, the type of referent achieved by the noun in the particular context is not one that would be contained in a standard lexical entry of the noun: no one thinks that the lexical entry of *ham sandwich* refers disjunctively to the food item or to customers who order it because of these uses, yet that the lexical entry of *hat* refers disjunctively to the clothing item or its stuff has been frequently advanced in the countability literature.

A second set of predictions concern when contextual variability of countability should NOT occur. There should be two sorts of gaps: (i) principled gaps, where a secondary use of the noun is unavailable since the resultant shift in meaning would conflict with the intrinsic meaning of the noun, and (ii) idiosyncratic gaps, where a secondary use of the noun could conceivably occur, but does not. Both of these gaps are found.

Despite various claims that all nouns have a countable and noncountable use, there are many types of nouns for which this appears impossible, which I term INFLEXIBLE NOUNS. For instance, nouns that designate a unit do not admit grinding: Galmiche (1989:68) observes the unacceptability in French of **du kilo*, **de la catégorie*, and **du chapitre*, which also holds for the English counterparts, **much kilo*, **much category*, or **much chapter*. Similarly, nouns designating two-dimensional shapes such as *triangle* possess only a countable use and are incompatible with grinding contexts.[34] While such restrictions would be puzzling under a view upon which the grammatical nature of nouns ensures equal access to countable and noncountable uses, they follow naturally from a meaning-based approach to countability: the descriptive content of nouns such as *kilometer* restricts their interpretation to conventionalized units of measurement, which is simply incompatible with referring to nondiscrete stuff.

Idiosyncratic gaps in contextual variability are well attested, even for English, which is quite permissive. For instance, grinding from an animal to its foodstuff is well attested, but fails from a fruit to its juice, as in 30. Many other such exceptions are given in Ostler & Atkins 1991 and Nunberg & Zaenen 1992.

(30) ??I enjoyed a glass of orange with my breakfast. (Nunberg 2004:352)

The availability of contextual variability also varies crosslinguistically. While some languages may have certain contextually licensed uses of bare nouns, these may differ from those in English. Nunberg (2004), citing personal communication with Jerrold Sadock, discusses the impossibility in Greenlandic Eskimo of canonical grinding from animals to their meat, but the presence of grinding from trees to their wood. (See also Cheng et al. 2008 for the failure of grinding in Mandarin.) Similar variation exists in the availability of packaging: Wiese and Maling (2005) demonstrate that even in closely related Germanic languages such as English, German, and Icelandic, the availability and grammatical strategies for packaging may differ substantively. The view that emerges of contextual variability is that cases of true polysemy fall on one extreme, and extralexical cases that succeed due to pragmatic reasoning fall at the other extreme.

[34] See Grimm 2012c and references therein for further discussion of inflexible nouns.

Between these two poles, there are many semi-productive relations between countable and noncountable uses that may hold, such as that a word referring to an animal may also refer to its meat in English, and less often fur. In addition, there is a large amount of conventionalization (see discussion in Ostler & Atkins 1991, Nunberg & Zaenen 1992). Yet here too the broad view on countability couched in individuation makes predictions that more structure-based accounts will not. The way in which speakers interact with things in the world is a predictor of their nominal descriptions' contextual-variability statuses. When we interact with a given type of thing only as discrete individual objects, such as for *hat*, then access to a meaning of, for instance, stuff is far less likely and is accordingly less likely to be stored in a lexical entry. Yet when we interact with some type of thing in the world both as discrete individual pieces and as nondiscrete portions, as in the case of string, then the noun describing this type of thing is more likely to refer to both discrete and nondiscrete quantities. Such a correlation is obvious, yet not tractable in theories of countability unless they incorporate how speakers interact with objects in the world.

5.2. Regularity and variation in countability classification: lexical and crosslinguistic. I now return to considering the puzzle, brought up in §1, of lexical variation (*foliage* versus *leaves*) and crosslinguistic variation (*hair* versus French *cheveux*) in countability classification. These doublets lead to an apparently paradoxical situation: how could two words that refer to the same things in the world belong to two different countability categories?

As observed by Joosten (2003), considering the relation between things and the world and their countability classification as based in objective ontological properties leaves unexplained these types of variation; however, considering the relation as arbitrary leaves unexplained the regularity witnessed in countability systems. Basing the relation between things and the world and their countability classification in individuation speaks to both the regularity and variability found.

Broad typological tendencies, such as substances being typically noncountable or animate entities being typically countable, are grounded in typical manners in which speech communities individuate and interact with such things. Another typological tendency, as seen in §2, is that entities in the world that come as aggregates are much more variable in how they are described within and across languages. This variation is, however, regular and also finds an explanation in an individuation-based account: it is precisely for these sorts of things that there is regularly a choice of how to individuate them and, consequently, how to describe them, as a single entity or as collections of entities.

As pointed out in §1, the puzzle of variability of countability classification arises only when under the assumption that a semantic theory of countability must be a direct, one-to-one relation between entities in the world and countability categories. It is clear that those making these arguments have such a theory in mind. For instance, as Chierchia (2010:151) puts it while discussing the crosslinguistic variation of the countability of hair in English and Italian (*capelli*, pl.): '*Hair*, used to refer to what grows on our head, seems to be mass in English, and count in Italian. Yet clearly we are referring to the same stuff. Your hair doesn't change, as we change language'.[35] This argument,

[35] Although clearly the physical hair would not change if the language changes, the description of hair may change as we change language, and this may have linguistic effects. Barner and Snedeker (2005:57, n. 7) found that French speakers 'judged six tiny pieces of hair to be "plus de cheveux" (i.e. more hairs) than two long ones, while English speakers picked the two long pieces as "more hair" '.

then, assumes that if countability classification is semantic, it should follow from the nature of the entity in the world that is referred to. In other words, this argument presupposes that objective ontological properties, such as the individual entity/stuff distinction, are the only relevant candidates that could influence nominal countability.

These arguments are effective against such a view of countability based in objective ontological properties, but do not engage the explanation of countability based in individuation properties. Most work on countability that invokes the notion of individuation clearly distinguishes entities in the world from CONSTRUALS of those entities (see McCawley 1975, Mufwene 1984, Bloom 1990, Wisniewski et al. 2003, inter alia). Thus, an individuation account entirely agrees with those who think that the explanatory burden should not be placed on the objective ontological properties of entities in the external world. Instead, the burden is placed on construals of entities in the world.

In a given situation, such as the sentences in 31, *leaves* and *foliage* may be interchangeable.

(31) a. The leaves have turned red.
b. The foliage has turned red.

Yet this apparent synonymy does not hold in general—the different nouns have different descriptive content, which lexicographers are careful to note. The dictionary entries for *leaf* and *foliage* from the *Oxford English Dictionary* (*OED*) are given in 32 and 33, respectively.

(32) *leaf*: 'An expanded organ of a plant, produced laterally from a stem or branch, or springing from its root; one of the parts of a plant which collectively constitute its foliage.' (*OED*)

(33) *foliage*: 'The leaves (of a plant or tree) collectively; leafage.' (*OED*)

These different descriptions are in turn associated with different individuation properties. For *leaves*, the focus is on the individual leaves, while for *foliage*, the focus is on the collectivity of leaves and other associated vegetation.

If within a language some entity may be described differently, it should be unsurprising that across different languages a given entity may be described differently, resulting in different countability classifications. Thus, mismatches that Chierchia brings up between *hair* and *capelli* (It.) should not be disturbing—although hair does not change as we change language, how hair is described may well change as we change language. This change, however, does not force one to conclude that no meaning-based categorization is in effect in countability classification.

Even when the descriptive content of cognates is nearly identical, the words in the different languages may have different countability classifications simply because the grammatical number systems in which they are found are structured differently. For instance, while English resorts to two different lexicalizations to convey the distinction between a collection of leaves and individual leaves, in languages with different morphological possibilities, these interpretations may be formally related. Thus, in discussing the Welsh grammatical number system, King (2003:67) states that '[t]he true relationship between collective and unit nouns is particularly clear in the English translations for some of the above pairs. *Foliage*, for example, is a very close approximation to the actual sense of *dail* [the collective form of "leaf"], conveying as it does the idea of "leaves as one homogeneous body" '. Thus, while English and Welsh presumably give voice to essentially the same meanings of collection of leaves, leaf, and plurality of leaves, because of the different structure of their grammatical number systems, these terms come out with different countability classifications in the different languages.

The scale of individuation makes clear predictions as to which portions of the lexicon will have a high degree of crosslinguistic agreement, namely the poles of the scale, and which sections of the lexicon will display less agreement, namely the middle region. This is exactly what one sees: although closely related languages like English and French may agree on nouns that correspond to the two poles of the scale, they differ on lexical items that fall somewhere in the middle zone. Thus, the fact that there are crosslinguistic mismatches does not invalidate the view of meaning-based categorization underlying countability classification; rather, these mismatches indicate for which entity types we must achieve a more nuanced understanding of the principles of meaning-based categorization underlying countability classification.

6. Conclusion. I have argued that grammatical number categorization reflects different degrees of individuation associated with nominal descriptions. This claim contains two parts. First, countability is not a binary distinction, although some languages may have only two primary grammatical number categories, namely countable and noncountable. Second, a noun's countability status is not purely a grammatical fact, but is based in individuation properties associated with the entity being described. Approached from this perspective, the putative paradoxes confronting a semantic grounding of countability resolve themselves. Variation among grammatical number systems, rather than being an obstacle for meaning-based categorization, points to regions of the lexicon that can add to our understanding of countability and sharpen our theories of word meaning and meaning-based categorization.

The framework presented here provides a broad perspective on countability and leaves open many paths for future investigation. Even within the English lexicon, much remains to be explored: a systematic study of the countability behavior of, for instance, abstract nouns or of more specific entity types that have come up in the discussion, such as 'paired entities' or 'types of people', will surely reveal more distinctions that must be accommodated. More typological work will undoubtedly lead to a more nuanced understanding of the scale of individuation. For instance, a balanced and thorough examination of the relation between grammatical number coding and entity types would clearly be a valuable contribution. Finally, further experimental work on the foundations of individuation may uncover still other factors interacting with the realization of grammatical number.

REFERENCES

Acquaviva, Paolo. 2008. *Lexical plurals: A morphosemantic approach*. Oxford: Oxford University Press.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21.435–83. DOI: 10.1023/A:1024109008573.

Allan, Keith. 1980. Nouns and countability. *Language* 56.541–67. DOI: 10.2307/414449.

Awbery, Gwenllian Mair. 2009. Pluralization in Welsh, and the need for phonologically null suffixes. *Journal of Celtic Linguistics* 13.1–20.

Baayen, R. Harald; Cristina Burani; and Robert Schreuder. 1997. Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. *Yearbook of Morphology* 1996.13–33. DOI: 10.1007/978-94-017-3718-0_3.

Baayen, R. Harald; Richard Piepenbrock; and Leon Gulickers. 1996. *CELEX2* [CD-ROM]. LDC96L14. Philadelphia: Linguistic Data Consortium.

Bale, Alan C., and David Barner. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26.217–52. DOI: 10.1093/jos/ffp003.

Barner, David, and Jesse Snedeker. 2005. Quantity judgments and individuation: Evidence that mass nouns count. *Cognition* 97.41–66. DOI: 10.1016/j.cognition.2004.06.009.

BLOOM, PAUL. 1990. *Semantic structure and language development*. Cambridge, MA: MIT dissertation.
BLOOM, PAUL. 1994. Possible names: The role of syntax-semantics mappings in the acquisition of nominals. *Lingua* 92.297–329. DOI: 10.1016/0024-3841(94)90345-X.
BLOOMFIELD, LEONARD. 1933. *Language*. New York: Henry Holt and Co.
BODOMO, ADAMS. 1997. *The structure of Dagaare*. Stanford, CA: CSLI Publications.
BODOMO, ADAMS. 2000. *Dagaare*. (Languages of the world/materials 165.) Munich: LINCOM Europa.
BORER, HAGIT. 2005. *Structuring sense, vol. 1: In name only*. Oxford: Oxford University Press.
BORG, ALBERT J. 1980. An aspectual distinction in verbs and nouns in Maltese. *Archivum Linguisticum* 11.107–17.
BRENNAN, ANDREW. 1988. *Conditions of identity*. Oxford: Clarendon.
BRINTON, LAUREL J., and ELIZABETH CLOSS TRAUGOTT. 2005. *Lexicalization and language change*. (Research surveys in linguistics.) Cambridge: Cambridge University Press.
BROWN, CECIL H. 1999. *Lexical acculturation in Native American languages*. (Oxford studies in anthropological linguistics 20.) Oxford: Oxford University Press.
BUNT, HARRY C. 1985. *Mass terms and model-theoretic semantics*. Cambridge: Cambridge University Press.
CHENG, LISA L. S.; JENNY DOETJES; and RINT SYBESMA. 2008. How universal is the universal grinder? *Linguistics in the Netherlands* 2008.50–62. DOI: 10.1075/avt.25.08che.
CHIERCHIA, GENNARO. 1998. Plurality of mass nouns and the notion of 'semantic parameter'. *Events and grammar*, ed. by Susan Rothstein, 53–104. Dordrecht: Kluwer.
CHIERCHIA, GENNARO. 2010. Language, thought, and reality after Chomsky. *The Chomsky notebook*, ed. by Julie Franck and Jean Bricmont, 142–69. New York: Columbia University Press.
CONTINI-MORAVA, ELLEN. 2000. Noun class as number in Swahili. *Between grammar and lexicon*, ed. by Ellen Contini-Morava and Yishai Tobin, 3–30. Amsterdam: John Benjamins.
CORBETT, GREVILLE G. 1996. Minor number and the plurality split. *Rivista di Linguistica* 8.101–22. Online: http://linguistica.sns.it/RdL/8.1/Corbett.pdf.
DAVIDSON, DONALD. 1973. On the very idea of a conceptual scheme. *Proceedings and Addresses of the American Philosophical Association* 47.5–20. DOI: 10.2307/3129898.
DENNY, J. PETER. 1976. What are noun classifiers good for? *Chicago Linguistic Society* 12. 122–32.
DIMMENDAAL, GERRIT J. 1983. *The Turkana language*. Dordrecht: Foris.
DIMMENDAAL, GERRIT J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42.214–61. Online: http://www.jstor.org/stable/30028548.
DORON, EDIT, and ANA MÜLLER. 2013. The cognitive basis of the mass-count distinction: Evidence from bare nouns. *Syntax and semantics, vol. 39: Crosslinguistic studies on noun phrase structure and reference*, ed. by Patricia Cabredo Hofherr and Anne Zribi-Hertz, 73–101. Leiden: Brill. DOI: 10.1163/9789004261440_005.
DOWTY, DAVID. 1991. Thematic proto-roles and argument selection. *Language* 67.547–619. DOI: 10.2307/415037.
FALKUM, INGRID LOSSIUS. 2010. Systematic polysemy and the count-mass distinction. *University College London Working Papers in Linguistics* 22.16–40.
FARGETTI, C. M. 2001. *Estudo fonológico e morfossintático da língua Juruna*. São Paulo: University of Campinas dissertation.
GALMICHE, MICHEL. 1989. Massif/comptable: De l'un à l'autre et inversement. *Termes massifs et termes comptables* (Recherches linguistiques 13), ed. by Jean David and Georges Kleiber, 63–77. Paris: Klincksieck.
GENTNER, DEDRE, and LERA BORODITSKY. 2001. Individuation, relational relativity and early word learning. *Language acquisition and conceptual development*, ed. by Melissa Bowerman and Stephen Levinson, 215–56. Cambridge: Cambridge University Press.
GIL, DAVID. 1996. Maltese 'collective' nouns: A typological perspective. *Rivista di Linguistica* 8.53–87. Online: http://linguistica.sns.it/RdL/8.1/Gil-collective.pdf.
GILLON, BRENDAN S. 1992. Towards a common semantics for English count and mass nouns. *Linguistics and Philosophy* 15.597–639. Online: http://www.jstor.org/stable/25001491.

Gleason, Henry A. 1965. *Linguistics and English grammar*. New York: Holt Rinehart and Winston.
Greenberg, Joseph. 1966. *Language universals*. The Hague: Mouton.
Grimm, Scott. 2012a. Degrees of countability: A mereotopological approach to the mass/count distinction. *Proceedings of Semantics and Linguistic Theory (SALT)* 22.584–603. DOI: 10.3765/salt.v22i0.2633.
Grimm, Scott. 2012b. Individuation and inverse number marking in Dagaare. *Count and mass across languages*, ed. by Diane Massam, 75–98. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199654277.003.0005.
Grimm, Scott. 2012c. *Number and individuation*. Stanford, CA: Stanford University dissertation.
Grimm, Scott, and Beth Levin. 2017. Artifact nouns: Reference and countability. *North East Linguistic Society (NELS)* 47.55–64.
Harbour, Daniel. 2008. *Morphosemantic number: From Kiowa noun classes to UG number features*. Dordrecht: Springer.
Haspelmath, Martin. 2005. Occurrence of nominal plurality. *The world atlas of language structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 142–45. Oxford: Oxford University Press.
Haspelmath, Martin. 2008. Creating economical morphosyntactic patterns in language change. *Linguistic universals and language change*, ed. by Jeff Good, 185–214. Oxford: Oxford University Press.
Horn, Laurence. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
Imai, Mutsumi, and Dedre Gentner. 1997. A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition* 62.169–200. DOI: 10.1016/S0010-0277(96)00784-6.
Jakobson, Roman. 1971 [1957]. Shifters, verbal categories, and the Russian verb. *Selected writings, vol. 2: Word and language*, ed. by Roman Jakobson, 386–92. The Hague: Mouton.
Jones, Morris, and Alan R. Thomas. 1977. *The Welsh language: Studies in its syntax and semantics*. Cardiff: University of Wales Press for the Schools Council.
Joosten, Frank. 2003. Accounts of the count-mass distinction: A critical survey. *Lingvisticæ Investigationes* 26.159–73. DOI: 10.1075/li.26.1.11joo.
Joosten, Frank; Gert De Sutter; Denis Drieghe; Stef Grondelaers; Robert J. Hartsuiker; and Dirk Speelman. 2007. Dutch collective nouns and conceptual profiling. *Linguistics* 45.85–132. DOI: 10.1515/LING.2007.003.
Katamba, Francis. 2003. Bantu nominal morphology. *The Bantu languages*, ed. by Derek Nurse and Gérard Philippson, 103–20. New York: Routledge.
King, Gareth. 2003. *Modern Welsh: A comprehensive grammar*. New York: Routledge.
Kiparsky, Paul, and Judith Tonhauser. 2012. Semantics of inflection. *Semantics: An international handbook of natural language meaning*, ed. by Claudia Maienborn, Klaus von Heusinger, and Paul Portner, 2070–97. Berlin: De Gruyter Mouton.
Koontz-Garboden, Andrew. 2007. *States, change of state, and the monotonicity hypothesis*. Stanford, CA: Stanford University dissertation.
Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. *Lexical matters*, ed. by Ivan Sag and Anna Szabolcsi, 29–53. Stanford, CA: CSLI Publications.
Landau, Barbara; Linda B. Smith; and Susan S. Jones. 1988. The importance of shape in early lexical learning. *Cognitive Development* 3.299–321. DOI: 10.1016/0885-2014(88)90014-7.
Landman, Fred. 1989. Groups, I. *Linguistics and Philosophy* 12.559–605. Online: http://www.jstor.org/stable/25001360.
Landman, Fred. 2011. Count nouns—mass nouns, neat nouns—mess nouns. *The Baltic International Yearbook of Cognition, Logic and Communication* 6. DOI: 10.4148/biyclc.v6i0.1579.
Lima, Suzi. 2014. *The grammar of individuation and counting*. Amherst: University of Massachusetts Amherst dissertation.
Lucy, John A. 1992. *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge, MA: Cambridge University Press.

MAHO, JOUNI FILIP. 1999. *A comparative study of Bantu noun classes*. Gothenburg: Acta Universitatis Gothoburgensis.
MATHIEU, ÉRIC. 2012. Flavors of division. *Linguistic Inquiry* 43.650–79. DOI: 10.1162/ling_a_00110.
MCCAWLEY, JAMES D. 1975. Lexicography and the count-mass distinction. *Berkeley Linguistics Society* 1.314–21. DOI: 10.3765/bls.v1i0.2335.
MIDDLETON, ERICA L.; EDWARD J. WISNIEWSKI; KELLY A. TRINDEL; and MUTSUMI IMAI. 2004. Separating the chaff from the oats: Evidence for a conceptual distinction between count noun and mass noun aggregates. *Journal of Memory and Language* 50.371–94. DOI: 10.1016/j.jml.2004.02.005.
MIFSUD, MANWEL. 1996. The collective in Maltese. *Rivista di Linguistica* 8.29–51. Online: http://linguistica.sns.it/RdL/8.1/Mifsud.pdf.
MOURELATOS, ALEXANDER P. D. 1978. Events, processes, and states. *Linguistics and Philosophy* 2.415–34. DOI: 10.1007/BF00149015.
MUFWENE, SALIKOKO S. 1980. Number, countability and markedness in Lingala LI-/MA- noun class. *Linguistics* 18.1019–52. DOI: 10.1515/ling.1980.18.11-12.1019.
MUFWENE, SALIKOKO S. 1984. The count/mass distinction and the English lexicon. *Chicago Linguistic Society (Parasession on lexical semantics)* 20(2).200–221.
NICHOLS, LYNN. 2008. Lexical semantic constraints on noun roots and noun borrowability. *Studies in Language* 32.683–700. DOI: 10.1075/sl.32.3.10nic.
NUNBERG, GEOFFREY. 1979. The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy* 3.143–84. Online: http://www.jstor.org/stable/25001016.
NUNBERG, GEOFFREY. 2004. Pragmatics of deferred interpretation. *The handbook of pragmatics*, ed. by Laurence R. Horn and Gregory Ward, 343–64. Oxford: Blackwell.
NUNBERG, GEOFFREY, and ANNIE ZAENEN. 1992. Systematic polysemy in lexicology and lexicography. *Euralex 1992, part 2*, 387–96. Online: http://euralex.org/publications/systematic-polysemy-in-lexicology-and-lexicography/.
NURMIO, SILVA. 2016. Collective nouns in Welsh: A noun category or a plural allomorph? *Transactions of the Philological Society* 115.1–21. DOI: 10.1111/1467-968X.12086.
OHTA, ITARU. 1989. *A classified vocabulary of the Turkana in Northwestern Kenya*. (African study monographs.) Kyoto: The Center for African Area Studies at Kyoto University.
OSTLER, NICHOLAS, and B. T. S. ATKINS. 1991. Predictable meaning shift: Some linguistic properties of lexical implication rules. *Lexical semantics and knowledge representation: SIGLEX 1991*, ed. by James Pustejovsky and Sabine Bergler, 87–100. Dordrecht: Springer. DOI: 10.1007/3-540-55801-2_29.
PALMER, FRANK R. 1971. *Grammar*. Harmondsworth: Penguin.
PELLETIER, FRANCIS JEFFRY. 1975. Non-singular reference: Some preliminaries. *Philosophia* 5.451–65.
PELLETIER, FRANCIS JEFFRY, and LENHART K. SCHUBERT. 2004. Mass expressions. *Handbook of philosophical logic*, vol. 10, ed. by Dov Gabbay and Franz Guenthner, 249–335. Dordrecht: Kluwer.
PRASADA, SANDEEP; KRAG FERENZ; and TODD HASKELL. 2002. Conceiving of entities as objects and as stuff. *Cognition* 83.141–65. DOI: 10.1016/S0010-0277(01)00173-1.
PRESS, IAN. 2009. Breton. *The Celtic languages*, ed. by Martin J. Ball and Nicole Müller, 427–87. New York: Routledge.
QUINE, W. V. O. 1960. *Word and object*. Cambridge, MA: MIT Press.
RAVID, DORIT, and LUBNA HAYEK. 2003. Learning about different ways of expressing number in the development of Palestinian Arabic. *First Language* 23.41–64. DOI: 10.1177/0142723703023001003.
ROTHSTEIN, SUSAN. 2010. Counting and the mass/count distinction. *Journal of Semantics* 27.343–97. DOI: 10.1093/jos/ffq007.
SAMUELSON, LARISSA K., and LINDA B. SMITH. 1999. Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition* 73.1–33. DOI: 10.1016/S0010-0277(99)00034-7.
SCHWARZSCHILD, ROGER. 2011. Stubborn distributivity, multiparticipant nouns and the count/mass distinction. *North East Linguistic Society (NELS)* 39(2).661–78.
SEIFART, FRANK. 2009. Towards a typology of unitization: Miraña noun classes compared to numeral classifiers and singulatives. Regensburg: University of Regensburg, MS.

Sharvy, Richard. 1978. Maybe English has no count nouns: Notes on Chinese semantics. *Studies in Language* 2.345–65. DOI: 10.1075/sl.2.3.04sha.
Smith-Stark, T. C. 1974. The plurality split. *Chicago Linguistic Society* 10.657–71.
Soja, Nancy N.; Susan Carey; and Elizabeth S. Spelke. 1991. Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition* 38.179–211. DOI: 10.1016/0010-0277(91)90051-5.
Stebbins, Tonya. 2007. Asymmetrical nominal number marking: A functional account. *STUF—Language Typology and Universals* 50.5–47. DOI: 10.1524/stuf.1997.50.1.5.
Stolz, Thomas. 2001. Singulative-collective: Natural morphology and stable cases in Welsh number inflexion on nouns. *STUF—Language Typology and Universals* 54.52–76. DOI: 10.1524/stuf.2001.54.1.52.
Sutcliffe, Edmund F. 1936. *A grammar of the Maltese languages*. Valletta: Progress.
Tiersma, Peter Meijes. 1982. Local and general markedness. *Language* 58.832–49. DOI: 10.2307/413959.
Ware, Robert X. 1975. Some bits and pieces. *Synthese* 31.379–93.
Watkins, Laurel. 1984. *A grammar of Kiowa*. Lincoln: University of Nebraska Press.
Wierzbicka, Anna. 1988. *The semantics of grammar*. Amsterdam: John Benjamins.
Wiese, Heike, and Joan Maling. 2005. *Beers*, *kaffi*, and *Schnaps*: Different grammatical options for restaurant talk coercions in three Germanic languages. *Journal of Germanic Linguistics* 17.1–38. DOI: 10.1017/S1470542705000012.
Wiggins, David. 1980. *Sameness and substance*. Oxford: Blackwell.
Wilhelm, Andrea. 2006. Count and mass nouns in Dëne Sųłiné. *West Coast Conference on Formal Linguistics (WCCFL)* 25.435–43. Online: http://www.lingref.com/cpp/wccfl/25/paper1477.pdf.
Wisniewski, Edward J.; Mutsumi Imai; and Lyman Casey. 1996. On the equivalence of superordinate concepts. *Cognition* 60.269–98. DOI: 10.1016/0010-0277(96)00707-X.
Wisniewski, Edward J.; Christopher A. Lamb; and Erica L. Middleton. 2003. On the conceptual basis for the count and mass noun distinction. *Language and Cognitive Processes* 18.583–624. DOI: 10.1080/01690960344000044.
Zwicky, Arnold. 2001. Counting chad. Handout from Stanford Semantics Fest, March 16, 2001.

[scott.grimm@rochester.edu]

[Received 1 April 2016;
revision invited 15 June 2016;
revision received 24 May 2017;
accepted pending revisions 22 September 2017;
revision received 22 November 2017;
accepted 13 December 2017]